THESE

présentée à

L'UNIVERSITE DE PROVENCE

AIX - MARSEILLE I

pour obtenir

LE GRADE DE DOCTEUR ÈS-SCIENCES PHYSIQUES

par

François SANTON

CONTRIBUTION AU PROBLEME DE L'INTELLIGIBILITE DE LA PAROLE DANS LES SALLES

Soutenue le 20 décembre 1974 devant la Commission d'Examen

MM. R. MERIGOUX — Président

P. CASAL
M. JESSEL
J-C. RISSET
R. SIESTRUNCK
— Examinateurs

THESE

présentée à

L'UNIVERSITE DE PROVENCE

AIX - MARSEILLE I

pour obtenir

LE GRADE DE DOCTEUR ÈS-SCIENCES PHYSIQUES

par

François SANTON

CONTRIBUTION AU PROBLEME DE L'INTELLIGIBILITE DE LA PAROLE DANS LES SALLES

Soutenue le 20 décembre 1974 devant la Commission d'Examen

MM.	R. MERIGOUX	Président
	P. CASAL	Examinateurs
	M. JESSEL	
	J-C. RISSET	
	R. SIESTRUNCK	

UNIV.
LEIDEN
BIBL.

- Professeurs sans chaire -

M. BOREL Fernand	Techniques mathématiques de la physique
M. BOUDET Roger	Mathématiques
Mle CATZIGRAS France	Géologie historique
Mle CHANDEBOIS Rosine	Biologie animale
M. COMBET Serge	Chimie
Mme DELACHE Solange	Mathématiques
M. FARDOUX Gérard	Mathématiques
M. FEUGEAS Claude	Chimie
M. FOUGEREAU Michel	Biochimie
M. GILLET André	Mathématiques
M. GUERMEUR Robert	Physique
M. HALBWACHS Francis	Physique
M. HENNEQUIN Jacques	Electronique
M. MANUCEAU Jérôme	Mathématiques
M. MERDY Henri	Physique
Mle PICHENOT Mathilde	Biologie végétale
M. PUJOL Louis	Chimie
M. SARDA Louis	Biochimie
Mme SIRUGUE Madeleine	Physique
M. TACHOIRE Henri	Chimie

- Maîtres de conférences -

M. BELAICH Jean-Pierre	Microbiologie
M. BENARROCHE Marcel	Physique
M. BENLIAN David	Chimie de coordination
M. BREUNEVAL Jacques	Mathématiques
M. CASALOT André	Chimie
M. CRESP Jean	Biologie animale
M. DELAAGE Michel	Biochimie
M. DELEUIL Roger	Physique
Mle DETRAZ Jacqueline	Mathématiques
M. LAVAGNE André	Biologie végétale
M. MARTINET Bernard	Mathématiques
M. MOULINS Maurice	Biologie végétale
M. PANTALONI Jacques	Thermodynamique
M. PHILIP Jean	Géologie historique
M. PUJOL Rémy	Psycho-physiologie
M. ROSAY Jean-Pierre	Mathématiques
M. SOUBLIN Jean-Pierre	Mathématiques
M. STREIFF Roland	Chimie

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41547006_0006

UNIVERSITE DE PROVENCE
(Aix-Marseille I)
Place Victor Hugo
13331 MARSEILLE CEDEX 3

- Service du Personnel -
5. ML.M/MG.

PERSONNEL ENSEIGNANT - Situation au 1.10.1974

(U.E.R. Scientifiques)

- Professeurs titulaires -

M. BODIOU Georges	Mathématiques appliquées
M. BODOT Hubert	Chimie
M. CARPENI Georges	Chimie
M. CASAL Pierre	Mécanique rationnelle et appliquée
M. DESNUELLE Pierre	Chimie biologique
M. FEHRENBACH Charles	Astronomie
M. FOURES Léonce	Mathématiques
M. GIRARDIE Adrien	Zoologie
M. GOUVERNET Claude	Géologie appliquée
M. HERVE Jacques	Electronique
M. HUGON Maurice	Psycho-physiologie
M. JULG André	Chimie théorique
M. LAFFITTE Marc	Chimie
M. MALE Damien	Physique
M. MERIGOUX Roger	Physique industrielle

- Professeurs titulaires à titre personnel -

M. BIZOUARD Michel	Physique
M. BLANCHARD André	Mathématiques générales
Mme BLANCHARD Christiane	Mathématiques
M. BONNEAU Marcel	Mathématiques
M. CADILHAC Michel	Physique
M. DOUCET Yves	Physique
M. FRAISSE Roland	Mathématiques
Mme FURNESTIN M.Louise	Biologie animale
M. GUENOCHE Henri	Mathématiques
M. MANDELBROJT Jacques	Physique
M. SIMON Claude	Physiologie animale
M. SOURIAU Jean-Marie	Méthodes mathématiques de la physique
M. VISCONTI Antoine	Physique théorique

REMERCIEMENTS

Messieurs les Professeurs Casal, Mérigoux, Risset et Siestrunck se sont intéressés à ce mémoire, et ont accepté de le juger. Je les prie de croire à mes remerciements respectueux.

Au moment où s'achève mon travail sur un sujet très vaste et par là même toujours délicat, je voudrais également remercier tous ceux qui m'ont guidé par leurs remarques, ou aidé amicalement.

Monsieur Th. Vogel m'a accueilli dans son Institut et a longtemps dirigé mes recherches. Monsieur M. Jessel ne m'a pas ménagé ses conseils et ses encouragements. Qu'ils veuillent trouver ici l'expression de ma profonde gratitude.

Monsieur H. Juricic m'a initié au calcul sur ordinateur et m'a beaucoup aidé dans l'élaboration des techniques faisant l'objet du chapitre 4 : notre collaboration a d'ailleurs fait l'objet d'un article paru dans la revue "Acustica" (1).

Je ne saurais oublier Monsieur B. Bladier avec qui j'ai fait mes premières armes dans le domaine épineux de l'acoustique des salles.

Ce travail n'aurait pas été réalisable sans l'aide des techniciens du service d'Acoustique Générale. A Th. Angélini et J. Mège, je dois l'entière réalisation du magnétophone à 10 têtes sur lequel reposent toutes les mesures du chapitre 5. Cet appareil est un témoignage éloquent de leur ingéniosité et de leur habileté. A André Daumas, je dois une assistance et une aide constante dans toutes les mesures, et en particulier les fastidieuses séances de logatomes, où sa patience a été mise à rude épreuve. G. Delaruotte a également collaboré avec nous, ici et là.

La gentillesse et l'efficacité de notre documentaliste, Madame D. Cattiaux, m'a été très précieuse. Madame G. Didry a parfaitement assuré la dactylographie de ce texte.

Les chefs de service des équipes d'Acoustique Appliquée ont mis à ma disposition leurs installations : en particulier, Monsieur G. Corsain a hébergé le "Champ Sonore Synthétique" dans sa vaste salle anéchoïque.

Que tous trouvent ici l'expression de mes remerciements les plus sincères.

x
x x
x

(1) Images et rayons sonores dans le calcul numérique des échogrammes - Acustica, Vol. 28, n° 2, pp. 77-89, 1973.

RESUME

Le problème de l'intelligibilité de la parole dans les salles relève du schéma général d'un système de communications : le message émis par un locuteur est déformé par un canal de transmission, en l'occurence la salle considérée. Il est reçu par l'auditeur, qui tâche de reconnaître le message émis (Ch. 1 et 2).

Pour savoir quel est le degré de reconnaissance du message reçu, on procède traditionnellement à des mesures d'intelligibilité. On fait réciter par un locuteur des listes de phrases, de mots ou de logatomes, c'est-à-dire de particules monosyllabiques du langage envisagé. Le pourcentage d'éléments phonétiques correctement perçus donne un "coefficient d'intelligibilité" variable, toutes choses égales d'ailleurs, suivant le type d'éléments phonétiques envisagés et suivant le nombre total de phrases, mots ou logatomes utilisés : une mesure d'intelligibilité n'a donc pas de signification absolue, et ne peut avoir qu'une valeur relative.

Le but de tout travail sur l'intelligibilité de la parole dans les salles est de proposer une prévision d'intelligibilité, c'est-à-dire une méthode qui, à partir des paramètres physiques de la salle envisagée, permette d'obtenir un coefficient aussi proche que possible du coefficient d'intelligibilité. Depuis les travaux de Knudsen, plusieurs estimations ont été proposées. Les premières, et les plus sommaires, reposent sur la seule notion de durée de réverbération; d'autres sont des adaptations de l'indice d'articulation des téléphonistes, au cas de l'acoustique des salles. D'autres méthodes proposent l'emploi des techniques de corrélation. Les procédés les plus naturels, à notre avis, sont ceux basés sur les échogrammes, ou diagrammes d'échos,illustrés en particulier par Lochner et Burger, Niese, et Januska. A notre tour, nous avons choisi d'utiliser les échogrammes pour proposer une estimation de l'intelligibilité (Ch. 3)

Les dernières méthodes mentionnées ci-dessus reposent sur l'enregistrement d'échogrammes expérimentaux : une source adéquate émet une

impulsion acoustique très brève, et l'on recueille au récepteur la "réponse de la salle", c'est-à-dire le son direct suivi d'une multitude d'échos d'intensité et de retard variables. Il nous a semblé utile d'envisager un calcul qui permettrait de connaître à l'avance l'échogramme, en considérant uniquement les plans de la salle et le coefficient d'absorption des parois. Il est encore très difficile d'utiliser un calcul basé sur la diffraction, qui donne des résultats tangibles dans le seul cas d'une excitation sinusoïdale. Le problème de la "réponse transitoire" d'une salle est heureusement directement accessible si l'on se contente du seul point de vue de l'acoustique géométrique.

Nous avons développé dans ce sens deux techniques: l'une est basée sur les images de la source réfléchie une ou plusieurs fois dans les différentes parois de la salle : c'est la méthode des sources images. L'autre suppose le rayonnement de la source discrétisé en une multitude de rayons sonores qui atteignent le récepteur après un nombre de réflexions variables, ou au contraire sont complétement absorbés par un panneau de la salle considérée, et par le sol en particulier. Ces deux méthodes donnent des résultats non pas identiques, mais très proches : d'autre part, le fait de connaître la direction d'arrivée des différents échos ne présente pas de difficulté.

Nous avons effectué une vérification expérimentale des échogrammes dans le cas d'une salle réverbérante heptaédrique, ainsi que dans celui d'une salle de conférences de géométrie relativement compliquée. Aussi curieux que cela puisse paraître, une telle confrontation n'avait jamais été effectuée, les uns accordant tout leur crédit aux calculs, les autres s'estimant satisfaits des résultats expérimentaux. (Ch. 4)

Pour exploiter les échogrammes spatio-temporels ainsi obtenus, nous avons entrepris des mesures du seuil de perception d'un écho, c'est-à-dire du seuil de nuisance d'une réflexion. En effet, les estimations de Lochner, Niese ou Januska reposent sur l'évaluation globale d'une énergie utile (énergie du son direct et des premières réflexions qui lui sont associées) et sur celle d'une énergie nuisible (énergie des réflexions qui arrivent au-delà d'un certain délai). Ces estimations ont toujours fait abstraction de la direction des échos, et la définition des frontières qui délimitent l'énergie utile et l'énergie nuisible est sujette à caution.

C'est pourquoi nous avons effectué à l'aide d'un champ sonore synthétique de nombreuses mesures du seuil de perception d'un écho. A une source primaire d'intensité I_o , constituée par le son direct seul, ou encore le son direct et une ou plusieurs réflexions qui lui sont parfaitement intégrées, c'est-à-dire ne provoquent aucun effet d'écho, on juxtapose une réflexion de retard τ et d'intensité I . En faisant varier le niveau de I , on constate que le seuil de perception I_s de l'écho dépend du retard et des directions angulaires de la réflexion considérée. Dans le cas où I est inférieur à I_s , l'énergie nuisible est nulle; dans le cas contraire, l'énergie utile est $I_u = I_o + I_s$, l'énergie nuisible $I_n = I - I_s$.

Après avoir tracé un réseau de courbes expérimentales qui donnent les variations du coefficient d'intelligibilité en fonction du niveau de la parole, en prenant un bruit de masque pour paramètre, on a constaté que l'évaluation de I_u et I_n permet une meilleure estimation de l'intelligibilité que celle donnée par les méthodes globales.
En particulier, certains résultats des mesures d'intelligibilité éloignés des prévisions habituelles trouvent leur justification dans le rôle perturbateur d'échos de même direction que le son direct. (Ch. 5)

x x x

S O M M A I R E

1. - INTRODUCTION

2. - LE SIGNAL EMIS : LA PAROLE

2.1- Voyelles et consonnes
2.2.- Le rôle des transitions
2.3.- La parole et l'information
2.4.- Quelques caractéristiques de la parole

3. - L'INTELLIGIBILITE DE LA PAROLE DANS LES SALLES

3.1.- Notion d'intelligibilité et mesure
3.1.1.- Phrases, mots et logatomes
3.1.2.- Les auditeurs
3.1.3.- Les locuteurs
3.1.4.- Méthode choisie pour mes mesures

3.2.- Méthodes de prévision de l'intelligibilité de la parole dans les salles
3.2.1.- Méthode de Knudsen
3.2.2.- Méthode des indices d'articulation : Bolt, Korn et Janssens
3.2.3.- Méthodes de corrélation
3.2.4.- La "fonction de transfert de modulation" de Houtgast et Steeneken
3.2.5.- Méthodes basées sur les échogrammes
a) les échogrammes
b) la "netteté" de Thiele
c) travaux de Lochner et Burger
d) travaux de Niese
e) travaux de Januska
f) méthode de Kürer : "centre de gravité" d'un échogramme.

3.2.6.- Quelques autres travaux sur l'intelligibilité de la parole

3.3.- Conclusion

4. - CALCUL D'UN ECHOGRAMME DANS UNE SALLE

4.1.- Mise en équation du problème

4.2.- Résolution à l'aide du régime permanent et de la transformation de Fourier

4.3.- Procédés de calcul au moyen de l'acoustique géométrique

4.3.1.- Objet de l'étude

4.3.2.- Aspect énergétique

4.3.3.- Méthode des sources images

a) cas de la salle rectangulaire

b) cas d'une salle quelconque

4.3.4.- Méthode des rayons discrets

4.3.5.- Introduction de l'aspect temporel

a) champ d'application et erreur des deux méthodes

b) réponse à une impulsion rectangulaire

c) réponse à une excitation quelconque

d) intégration auditive d'un signal

e) remarque sur la réponse en fréquence

4.3.6.- Application numérique

a) salle rectangulaire

b) salle heptaédrique

c) efficacité des deux méthodes

4.3.7.- Vérification expérimentale

a) remarque sur le spectre de l'impulsion émise

b) remarque sur la forme de l'impulsion

c) montage expérimental

d) mesure des coefficients de réflexion des parois

e) salle heptaédrique

f) salle de conférences

4.4.- Conclusion

5. - MESURES ET PREVISIONS D'INTELLIGIBILITE

5.1.- Remarque sur l'intelligibilité en présence de bruit
5.2.- Mesures subjectives d'intelligibilité
5.3.- Prévisions objectives d'intelligibilité, à l'aide des échogrammes temporels
a) Méthode hybride
b) Evaluation de Thiele-Januska
c) Evaluation de Niese
d) Evaluation de Lochner et Burger
e) Conclusion
5.4.- Le champ sonore synthétique
5.5.- Mesure du seuil de perception des échos
5.5.1.- Montage utilisé et mode opératoire
a) Définition du seuil de perception des échos
b) Définition du délai critique
c) Comparaison des deux méthodes
d) Définition du seuil de perception absolu d'une réflexion
5.5.2.- Seuil de perception de l'écho, voie 2
a) Remarque sur la direction du son direct
b) Remarque sur l'intensité du son réfléchi
c) Cas d'une source primaire multidirectionnelle
5.5.3.- Seuil de perception de l'écho, voie 3
5.5.4.- Seuil de perception de l'écho, voie 4
a) Cas des voies 2 et 3 monodirectionnelles
b) Cas des voies 2 et 3 multidirectionnelles.
5.6.- Utilisation des courbes de seuil de perception des échos pour la prévision de l'intelligibilité.
5.6.1.- Généralités
5.6.2.- Mesures d'intelligibilité en présence d'un bruit de direction variable
a) Cas d'une source primaire monodirectionnelle
b) cas d'une source primaire multidirectionnelle

c) Intelligibilité en présence d'un bruit diffus

5.6.3.- Cas de l'écho, voie 1

5.6.4.- Cas de plusieurs échos

5.7.- Conclusion

6. - CONCLUSION GENERALE

x

x x

x

1.- I N T R O D U C T I O N

Le problème de l'intelligibilité de la parole dans les salles est particulièrement vaste. C'est une question bien caractéristique du domaine de l'acoustique, qui est une science diagonale, ou une interdiscipline, comme le faisait remarquer M. Jessel [1.1] .
Deux éléments fondamentaux apparaissent, lorsqu'on envisage le seul titre de ce travail : d'une part, l'intelligibilité de la parole, domaine très important, exploité de façon intensive par les phonéticiens et les spécialistes des télécommunications; d'autre part, l'acoustique des salles, responsable de la déformation apportée par la salle au signal acoustique émis.

Un problème est de savoir, quand on veut aborder ce domaine, quelle part réserver à l'intelligibilité, quelle part réserver à l'acoustique des salles. Il est possible de considérer qu'une salle ne transforme la parole qu'en lui ajoutant une "réverbération" analogue à un autre brouillage. On réduit alors le problème à celui de l'intelligibilité, d'une façon générale, c'est-à-dire à celui des communications à courtes ou grandes distances. En fait, étant donné l'importance primordiale de l'intelligibilité dans le domaine des télécommunications, c'est ce problème qui a été de loin le plus exploité. D'où la tendance naturelle de nombreux auteurs à ramener le cas de l'intelligibilité de la parole dans les salles à ce domaine mieux connu, et à considérer le problème des salles comme un simple "cas particulier" de celui du bruit de fond.
Cette démarche est, à beaucoup de points de vue très logique, et d'ailleurs, par la suite nous serons amenés à considérer effectivement une énergie utile et une énergie nuisible, intimement mêlées dans le signal reçu.

On définit ainsi implicitement les deux facteurs dont dépendra l'intelligibilité : le niveau du son utile, et celui du son nuisible qu'on peut remplacer par le rapport Signal/Bruit.
Cependant, il est un autre élément dont le rôle est loin d'être négligeable,

c'est celui de l'audition binaurale qui apporte au cas de l'intelligibilité de la parole dans les salles une originalité particulière.

Pour situer le problème dans toute son ampleur et sa diversité, nous reprendrons le schéma proposé par Shannon [1.2] , donné par la figure 1.1 dont nous proposons une version plus détaillée (fig. 1.2)

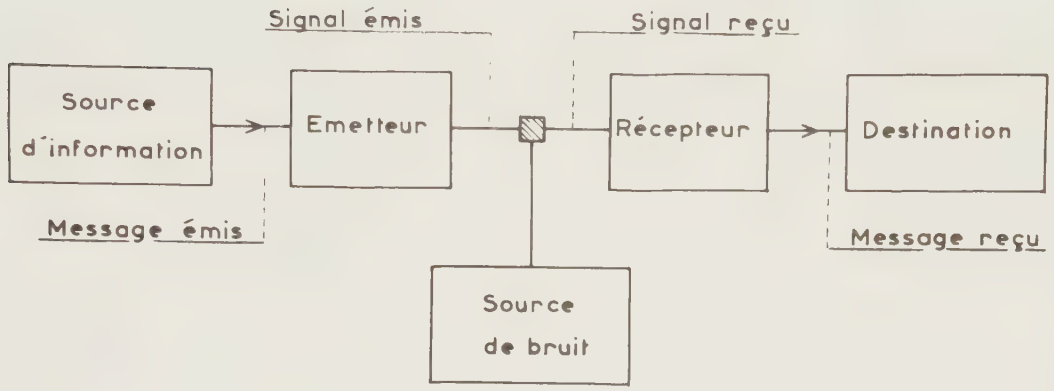


Fig. 1.1- Schéma général d'un système de communications - d'après SHANNON -

Une "source d'information", c'est-à-dire le cerveau du locuteur transmet un "message" à l'émetteur acoustique, c'est-à-dire aux organes de la phonation. Le message acoustique, la parole, est à son tour transmis par un canal constitué, dans le cas qui nous intéresse, par la salle. Le signal acoustique transformé par la salle est alors reçu par les deux oreilles de l'auditeur : par les passages successifs de l'oreille externe, de l'oreille moyenne et de l'oreille interne, ce signal arrive aux nerfs auditifs qui le transmettent au cerveau. A ce niveau s'effectue à proprement parler l'opération de reconnaissance qui permettra à l'auditeur de dire s'il a reconnu tel ou tel phonème, phonatome, ou tel mot.

Le problème relève donc de la phonétique, au niveau du signal acoustique émis : la structure particulière de la parole a été mise en

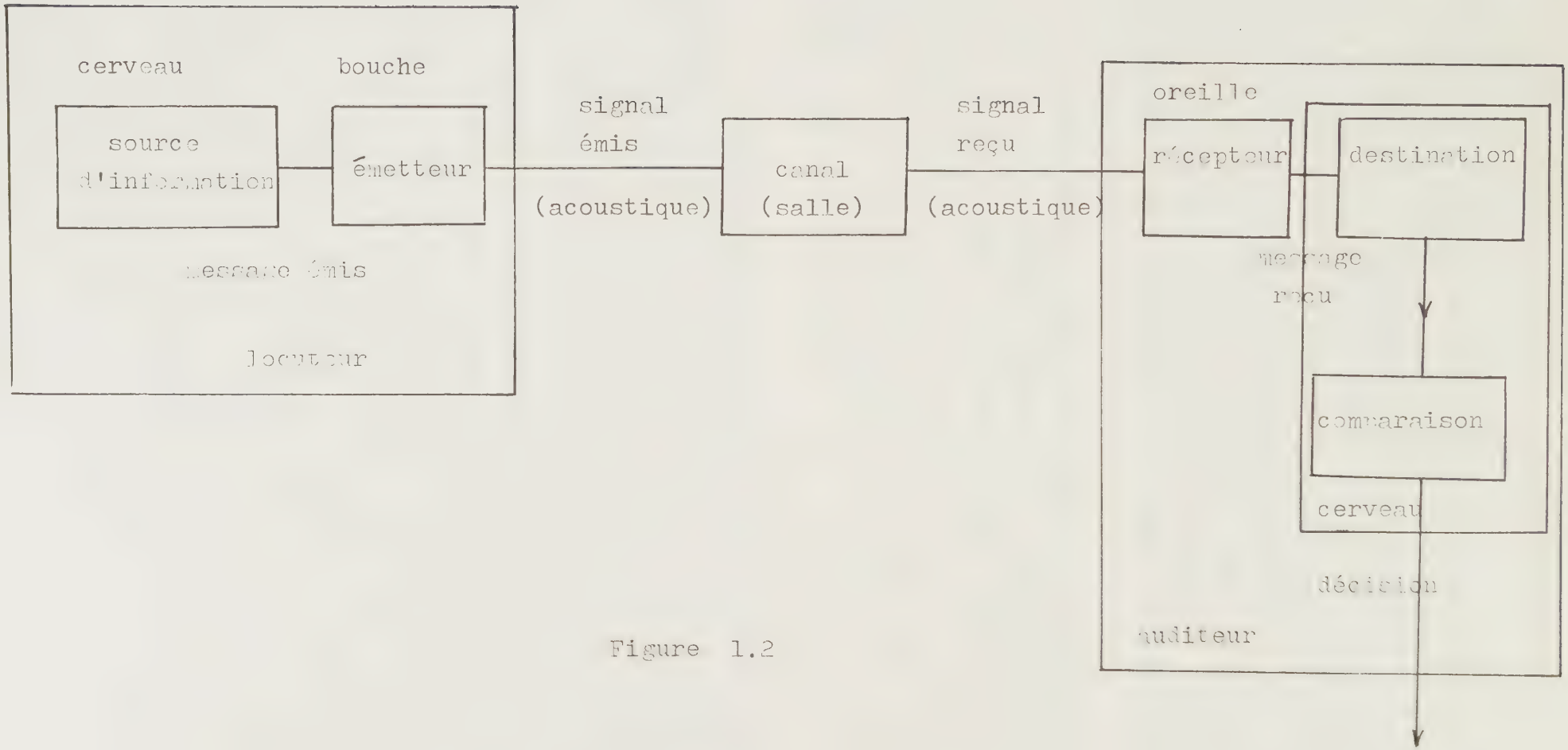

cerveau
bouche
source
d'information
émetteur
message émis
locuteur
signal
émis
(acoustique)
canal
(salle)
signal
reçu
(acoustique)
oreille
récepteur
destination
message
reçu
comparaison
cerveau
décision
auditeur

Figure 1.2

lumière par les phonéticiens. L'acoustique des salles a un rôle très important, puisque c'est elle qui va déterminer la transformation subie par la parole. Enfin, comme le résultat de la mesure d'intelligibilité sera dû au jugement de l'auditeur, le dernier stade du problème, dont l'importance n'est pas la moindre, est du domaine psychoacoustique : les caractéristiques de l'oreille jouent un rôle considérable, en particulier dans le cas de l'acoustique des salles où l'on a affaire aux "échos" dont l'effet sera très différent, suivant leur délai par rapport au son direct. D'autre part, le phénomène cérébral de "reconnaissance des formes" est très important, puisqu'il détermine la réaction de l'auditeur.

Notre étude se situe principalement au niveau de l'acoustique des salles. Il nous a semblé intéressant d'utiliser des méthodes de calcul numérique, déjà préconisées par certains auteurs, mais en leur donnant un développement qu'elles n'avaient pas connu jusqu'alors. Ces méthodes de calcul débouchent sur la "réponse transitoire d'une salle", ou la détermination de la répartition spatio-temporelle des échos. Ces calculs constituent la première partie de notre travail.

Cependant, un échogramme, aussi exact soit-il, ne permet pas de donner une estimation de l'intelligibilité, si l'on ne tient pas compte des caractéristiques de l'audition, et de l'audition binaurale en particulier. Car le but de toute recherche sur l'intelligibilité de la parole est de relier la "mesure directe", effectuée à l'aide l'éléments phonétiques prononcés par des locuteurs et reçus par des auditeurs, à une prévision qui sera faite en estimant les divers paramètres physiques d'une salle susceptibles de modifier le signal émis. De telles prévisions existent, mais même les meilleures ne tiennent pas compte d'un élément qui risque, à notre avis, d'avoir son importance dans certains cas : c'est l'audition binaurale.

Sans avoir la prétention d'expliquer quels sont les phénomènes qui peuvent justifier le processus de l'intégration auditive des échos, il nous a paru indispensable de faire des mesures qui mettent en évidence ces phénomènes : tel est le but de la deuxième partie de notre travail, basée sur l'exploitation d'un "champ sonore synthétique" qui simule les effets de la réverbération, et permet de modifier facilement tel ou tel paramètre.

Nous envisagerons, mais succintement, les processus de "reconnais-

sance des formes". Bien sûr, l'importance de ce phénomène est considérable, et même déterminante. Cependant, il ne faut pas se leurrer sur tel ou tel processus physique qui permettrait de l'appréhender facilement. Nous évoquerons les méthodes de comparaison optique des "formes sonagraphiques", ainsi que les méthodes de corrélation : certaines portes sont ouvertes, elles ne débouchent pas pour autant sur une solution automatique, et il convient d'être très prudent.

Il n'est pas de notre ressort de nous étendre outre mesure sur la structure physique du signal acoustique émis, c'est-à-dire la parole. Les travaux très nombreux des spécialistes de l'analyse et de la synthèse de la parole donnent des renseignements exhaustifs. Cependant, nous résumerons des caractéristiques principales de ce signal très particulier, qui est le matériau même et l'objet de la transmission.

x x x

B I B L I O G R A P H I E

1.1. JESSEL (M.).- Acoustique théorique et holophonie (Masson, 1973)

1.2. SHANNON (C.E.).- A mathematical theory of communications. (The Bell system technical journal, p.381, 1948)

x x x

2.- LE SIGNAL EMIS : LA PAROLE -

2.1.- Voyelles et consonnes.-

Le signal acoustique de la parole est constitué par un assemblage de sons, les uns ayant un caractère harmonique, ou encore musical, les autres étant des bruits de caractère typiquement transitoire, ou au contraire quasi-stationnaire. Dans le premier groupe, on aura tendance à classer les voyelles, dans le second, les consonnes.

Cette classification schématique est inexacte, car certaines consonnes sont sonores, c'est-à-dire sont des bruits combinés avec un ton laryngien, par exemple les consonnes g, v, z. En fait, on peut définir les consonnes comme étant caractérisées par un rétrécissement ou une fermeture momentanée du passage de l'air dans le conduit vocal. Les consonnes typiquement transitoires sont justement les occlusives telles que p, d, t, k... alors que d'autres peuvent être soutenues, comme les fricatives f,s,v, z, r, ch. La même distinction peut se retrouver dans le cas des voyelles, étant donné l'existence, dans la langue française entre autres, de voyelles à résonance nasale, telles que an, in, on, un, dont le caractère est partiellement transitoire.

En résumé, la structure acoustique de la parole est constituée par des sons harmoniques ou des bruits dont l'enveloppe du spectre, dans le domaine des fréquences, est caractéristique des différents phonèmes.
Les phonéticiens appelent "régions formantiques" les maximums locaux de ces spectres, en particulier dans le cas des voyelles, et ont ainsi défini un "triangle de formants" où l'on situe une voyelle en fonction des fréquences des deux premiers formants, comme le montre la figure 2.1.

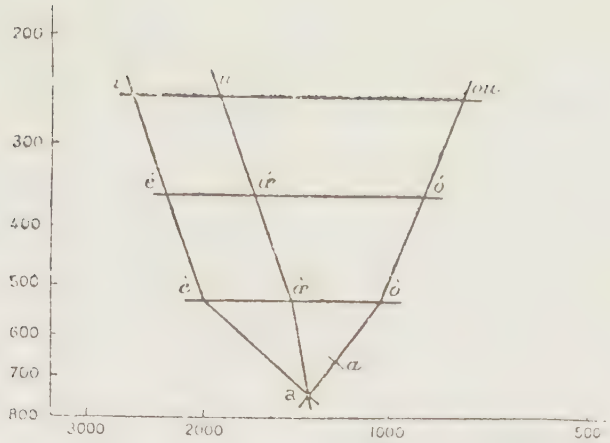


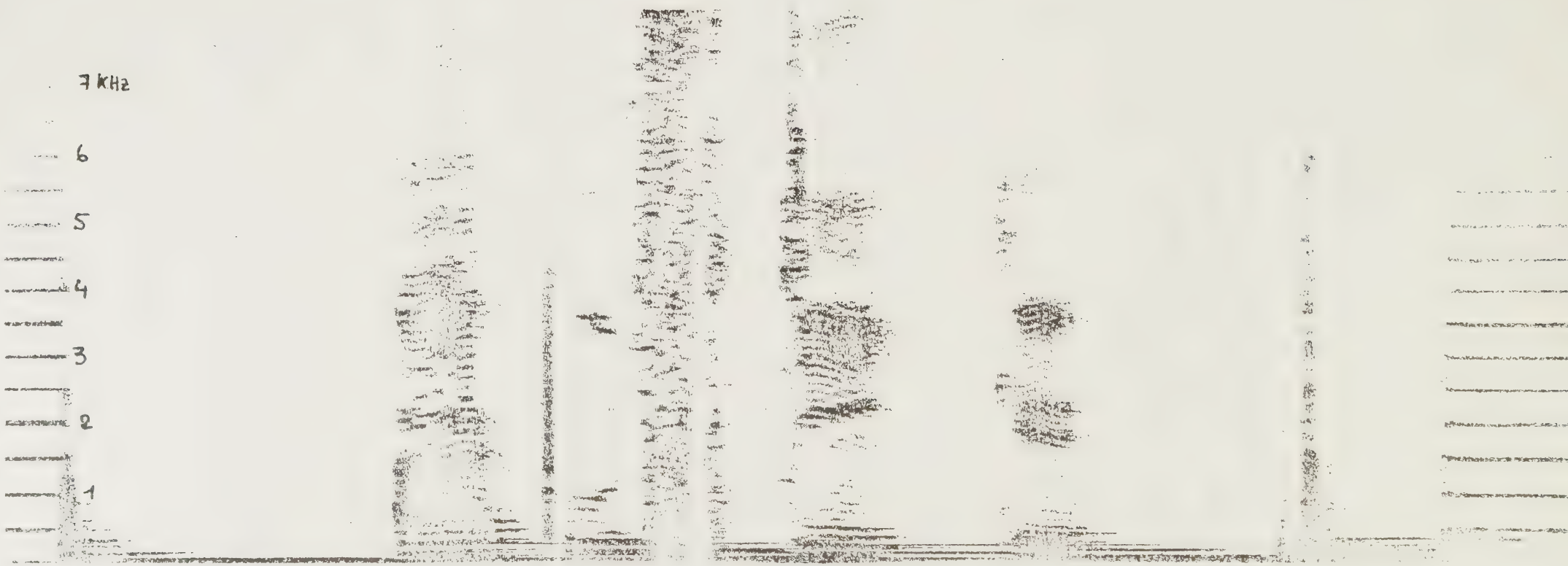
Fig. 2.1.- Schéma vocalique des voyelles françaises, d'après MALMBERG. En abscisse, la fréquence du second formant, en ordonnées celle du premier, en Herz également.-

2.2.- Rôle des transitoires.-

Dès les travaux de Delattre [2.1.] , en 1950, et l'utilisation du sonagraphe, puis d'un "play-back" qui permettait de lire les enregistrements du sonagraphe, on s'est aperçu de l'importance des transitions entre les différents phonèmes. Avec l'utilisation de l'Icophone, plus récemment construit au Laboratoire d'Acoustique Musicale de l'Université de Paris VI, LEIPP a mis en évidence le rôle des phonatomes [2.2.] , [2.3.] , assemblages insécables de phonèmes. Ainsi, le mot "Paris" ne sera pas décomposé en quatre phonèmes p, a, r et i, mais en trois phonatomes pa, ar, et ri.

A ce sujet, il faut insister sur la mine de renseignements donnés par le sonagraphe, cet enregistreur graphique fabriqué par Kay-Electric, puis perfectionné par la compagnie Bell avec le "Voice-Print",et qui donne une image particulièrement significative d'un message acoustique en général, et de la parole en particulier : les trois paramètres essentiels y apparaissent, le temps en abscisse, la fréquence en ordonnées, et l'intensité, à laquelle le noircissement est proportionnel (fig. 2.2.)

Fig 2.2. Courbe mélodique du mot "accoustique"
prononcé par une personne sourde.



2.3.- La parole et l'information.- [2.4]

On parle pour communiquer, cette constatation est d'une évidence banale. Mais on peut également remarquer que le musicien compose de la musique, que l'instrumentiste joue, et communique ainsi avec un auditeur. Toute combinatoire sonore est un moyen de communication, comme toute oeuvre d'art, toute expression élaborée ou non. Si l'on se limite aux communications purement acoustiques, on saisit immédiatement quelle différence sépare le problème de la parole et celui de la musique. Cela tient à la dualité de l'information contenue dans ces messages : une partie de l'information sera dite "sémantique", une autre"esthétique". La frontière entre ces deux types d'informations est assez floue, il faut en convenir. On peut dire que l'information sémantique de la parole est l'information "brute". Si l'on fait lire une phrase du langage courant à plusieurs locuteurs de sexe, d'âge, et de nationalités différentes, et ceci en plusieurs langues, un auditeur polyglotte trouvera dans le message reçu une information commune, indiquant, par exemple, qu'il fait beau aujourd'hui. Toutes les observations que pourra faire l'auditeur, relatives au caractère de la voix, à sa mélodie, à ses intonations, etc..., seront de l'ordre de l'information "esthétique" [2.6.] Il apparaît que la part de l'information esthétique véhiculée par la parole est assez importante. On peut estimer que cette information n'est pas indispensable, un speaker ou un autre pouvant par exemple annoncer telle nouvelle à la radio, sans que le résultat, pour l'auditeur, soit sensiblement altéré. Le problème de la compréhension du message, lié, ainsi que nous pourrons le mentionner plus loin, à celui de l'intelligibilité, dépend par conséquent de la perception de l'information "sémantique" de la parole.

On peut avoir une idée du rapport numérique de ces informations en comparant l'information phonétique à l'information totale.

Considérons un message constitué d'une suite de n éléments distincts, choisis parmi un ensemble de N éléments, chaque élément ayant une probabilité d'apparition p_i . On associe à ce message une quantité d'information :

$$H = -n \sum_{i=1}^{N} p_i \log_2 p_i$$

Si les messages sont équiprobables, $p_i = p = 1/N$ pour tout i, alors :

$$H = -n \log_2 p$$

H est exprimé en digits binaires (bits) . Dans le cas d'un message de bande passante ΔF et de dynamique D, et si l'on suppose suffisant un échantillonnage de fréquence $2\Delta F$, et la présence de D éléments différentiels de niveau équiprobable, on aboutit à la formule de Hartley :

$$H = 2\Delta F \log_2 D$$

Si $\Delta F = 1{,}5.\ 10^4$ Hz, ce qui représente la bande des fréquences audibles, et $D = 10^6$, soit 60 dB, on trouve un débit d'information de $4.\ 10^5$ bits par seconde. Cette information est l'information totale véhiculée par le message.

Considérons maintenant l'information liée aux phonèmes : supposons que leur nombre soit de l'ordre de $32 = 2^5$, et qu'un locuteur en prononce 10 par seconde . L'information phonétique sera alors :

$$H_1 = -10 \log_2 2^{-5} = 50 \text{ bits/s}$$

Il apparaît nettement que cette information, dont l'estimation est d'ailleurs assez grossière, les phonèmes n'étant pas équiprobables, est nettement inférieure à l'information totale véhiculée par la parole.

2.4.- Quelques caractéristiques de la parole.-

En regardant un enregistrement au sonagraphe, on distingue nettement que les voyelles et les consonnes ont des durées très différentes. En général, les voyelles ont une durée de l'ordre de 150 à 200 ms; au contraire, les consonnes ont une durée beaucoup plus réduite (de l'ordre

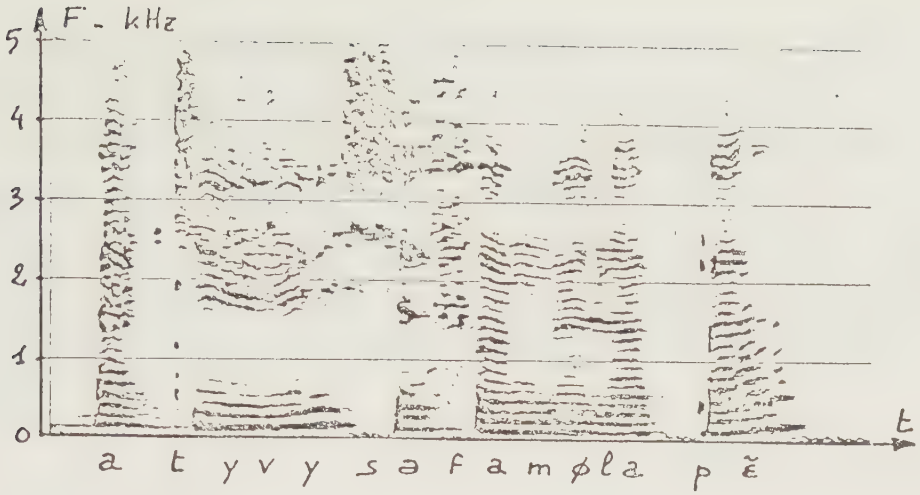
de 30 à 50 ms). La puissance des voyelles est, d'après Fletcher [2.5] comprise entre 250 μw, pour "ou" et 600 μw, pour "a". Si les consonnes voisées ont une puissance comparable à celle des voyelles, celle des consonnes occlusives en particulier est très faible, soit 10 μw pour p, t, d.

Il apparaît ainsi que les phonèmes les plus importants pour la compréhension, c'est-à-dire les consonnes, sont les plus brefs et les plus faibles. Au contraire, les moins importants, c'est-à-dire les voyelles, ont une énergie beaucoup plus grande, et une durée également supérieure. Des expériences ont montré qu'il était tout-à-fait possible de réduire la durée des voyelles sans entamer pour autant la compréhension du discours.

Un autre élément est à prendre en considération; si la bande passante de la parole s'étend de 100 à 8000 Hz, on peut réduire cette bande par filtrage à la zone 300-4000 Hz sans altérer la compréhension du discours. La possibilité de compression de la bande passante est dûe à la situation des régions formantiques dans la bande centrale du spectre de la parole.

Une autre caractéristique de la parole peut être également comprimée : il s'agit de la dynamique. Si elle peut atteindre 80 dB dans le cas de la voix criée, elle est réduite à 15 dB pour la voix chuchotée : or la voix chuchotée reste très compréhensible.

Il apparaît par conséquent que la parole contient une information sémantique dont l'importance est capitale, mais qui ne représente qu'une partie assez restreinte de l'information totale.On peut, par divers procédés, réduire l'information totale, tout en conservant l'information sémantique, c'est-à-dire celle qui s'adresse à la "compréhension" de l'auditeur : si l'information est comprimée, l'intelligibilité peut demeurer très élevée, et toute l'information sémantique sera conservée. Ceci dit, il peut paraître aventureux de présenter une "information squelettique" comme celle schématisée dans les dessins de type sonagraphique de Delattre ou de Leipp, comme l'information sémantique. (Fig. 2.3.)



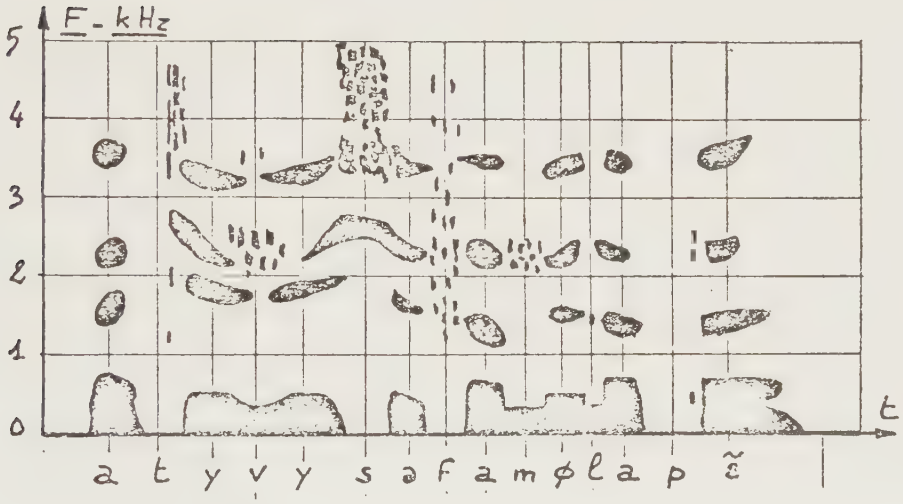


Fig 2.3. Sonagramme de la phrase "as tu vu ce fameux lapin", et squelette sémantique, d'après J.S. Lienard.

D'autre part, l'information esthétique ne peut pas être comprimée sans que la qualité n'en souffre. On sait que le matériel "haute fidélité" a justement pour objet de retransmettre la quasi-totalité du message à lire (disque, bande magnétique, etc...).
Au contraire, la qualité d'une information sémantique, c'est -à-dire la compréhension, peut résister à une compression de l'information : il importe maintenant de préciser la notion de compréhension, et de définir celle d'intelligibilité.

x x x

B I B L I O G R A P H I E

2.1. DELATTRE (P.C.) et alii.- Some experiments on the perception of synthetic speech sounds.- JASA, V24, n° 6, p.597-606, 1952.

2.2. LEIPP (E.).- Information sémantique et parole.- Bulletin du G.A.M., n° 22, juin 1966.

2.3. LEIPP (E.) et alii.- Structure physique et contenu sémantique de la parole.- Revue d'acoustique, n° 3-4, p.259,1968.

2.4. LIENARD (J.S.).- Analyse, synthèse et reconnaissance automatique de la parole.- Thèse de Doctorat d'Etat, Université de Paris VI, avril 1972.

2.5. FLETCHER (H.).- Speech and hearing in communications.- Van Nostrand, Ed, New-York, 1953.

2.6. MOLES (A.)- Théorie de l'information et perception esthétique.- Flammarion, Ed, 1958.

Nous avons également consulté les ouvrages suivants :

- FANT G.- Acoustic theory of speech production.
Mouton et Cie, Ed, La Haye, 1960.

- MALMBERG B.- La phonétique. - PUF.

- MOLES A et VALLANCIEN B.- Phonétique et phonation.
Masson, ed, Paris, 1966.

x x x

3.- L'INTELLIGIBILITE DE LA PAROLE DANS LES SALLES -

3.1.- La notion d'intelligibilité et les mesures.-

Le rôle d'un bon système de transmission est de permettre à l'auditeur de saisir entièrement l'information sémantique du discours prononcé par le locuteur : c'est à ce critère de qualité que nous nous arrêterons. Bien entendu, cette qualité sera différente de celle qu'on exigerait d'une salle de concert qui doit modifier le signal acoustique émis, en l'occurence une information esthétique, suivant certains critères eux aussi d'ordre esthétique, établis par l'expérience, ou le cas échéant demandés par le compositeur.

Bien "entendre" la parole, c'est bien la comprendre; l'expression ancienne est éloquente à ce sujet. La compréhension d'un discours, d'une communication ne dépend pas uniquement de la qualité du système de transmission, mais également de l'attention de l'auditeur, de ses capacités intellectuelles, etc... En fait, pour chiffrer la qualité d'une salle relative à la retransmission d'une communication parlée, il faut faire abstraction, autant que possible, de ces facteurs.

Dans la hiérarchie des perceptions qui entrent en jeu lors de l'audition de la parole, on peut considérer :

- 1) La perception des phonèmes ou des phonatomes -
- 2) La compréhension des mots -
- 3) La compréhension des phrases -

On conçoit facilement que, si telle conférence est jugée incompréhensible par les auditeurs, l'acoustique de la salle risque fort de n'être pas seule en cause : pour s'affranchir du problème du langage,ou du vocabulaire commun à l'auditeur et au locuteur, on a proposé l'utilisation de mots brefs et sans signification, pour éprouver la qualité d'un système

de transmission de la parole. La perception de ces "logatomes" relève de la notion de reconnaissance des formes, mais à un stade beaucoup plus élémentaire. Certains phonatomes étrangers au français seraient fatalement moins bien perçus par des auditeurs français que ceux qui y sont fréquents.

3.1.1.- Phrases, mots et logatomes.-

On peut effectuer une mesure d'intelligibilité en utilisant des groupes phonétiques courts, appelés "logatomes", suivant l'expression proposée par Stevenson en 1930. On définit ainsi la "netteté phonétique": c'est le rapport du nombre de groupes phonétiques correctement perçus par l'auditeur au nombre de groupes phonétiques émis. [3.1.] , [3.2.] . Cette netteté phonétique est souvent appelée "coefficient d'intelligibilité" ou plus brièvement "intelligibilité". Cette dénomination usuelle n'est peut-être pas très heureuse, car elle laisse entendre que l'intelligence jour un rôle dans la perception des phonèmes, ce qui semble à priori improbable. L'expression "coefficient de perception "serait peut-être plus appropriée. Les auteurs anglais ou américains précisent que la mesure a été faite sur des logatomes monosyllabiques, avec l'expression "syllabe intelligibility".

Aux logatomes, on peut préférer des mots qui seront eux aussi constitués en listes. On mesure ainsi un coefficient différent du précédent, et qui sera relatif à la compréhension des mots. Si ces mots font partie du langage courant, on peut estimer que l'intelligence saura suppléer aux déficiences éventuelles de la perception de tel phonème. Dans ce cas, le coefficient mesuré avec des mots sera supérieur à celui donné, dans les mêmes conditions acoustiques, par des logatomes.
Si, au contraire, un mot est inconnu, c'est l'inverse qui se produira : on sait que la perception correcte d'un nom propre, s'il n'est pas usuel ou porté par un homme illustre, est très difficile, et demande presque toujours une répétition. On peut également établir des listes de phrases simples, du langage usuel. Le coefficient d'intelligibilité mesuré sera alors plus élevé que celui mesuré avec des mots, pour des raisons analogues à celle évoquée plus haut. Mais le résultat serait opposé si l'on utilisait des phrases extraites de poésie moderne. (Fig. 3.1.).

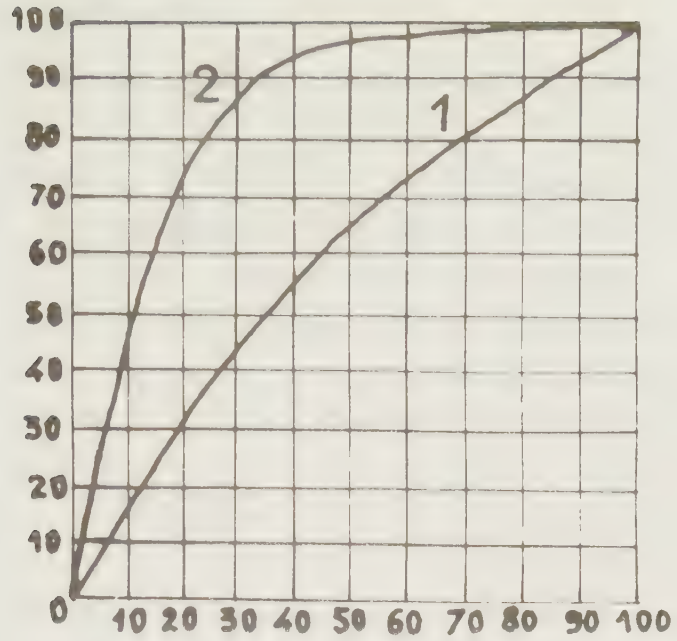


Fig. 3.1.- Intelligibilité en %, pour des mots(1) ou des phrases (2) en fonction de celle mesurée sur des logatomes, d'après Chavasse [2.3.]

Bien entendu, les listes constituées doivent présenter des phonèmes avec des fréquences correspondant à celles de la langue usuelle "moyenne". La constitution de ces listes relève par conséquent de la phonétique statistique, on peut la mener à bien par des moyens artisanaux, et un travail de bénédictin [2.4.] [3.3.] ; on préfère utiliser maintenant des moyens puissants, et établir ces listes à l'aide d'ordinateurs. [3.4.] De telles listes seront dites "phonétiquement équilibrées" ("phonetically-balanced", en anglais).

Que choisir entre les listes, de phrases, de mots ou de logatomes? A priori, il est impossible de justifier tel ou tel choix, si l'on ne tient pas compte des caractéristiques des autres instruments de mesure, c'est-à-dire des auditeurs, de leurs possibilités d'apprentissage, etc... Examinons maintenant quels sont les problèmes posés par le choix des locuteurs et celui des auditeurs.

3.1.2.- Les auditeurs -

Le rôle des auditeurs est le suivant : ils reçoivent les logatomes émis dans la salle, ils les perçoivent correctement ou non : pour connaître le résultat, la méthode la plus rudimentaire et la plus utilisée est de demander aux auditeurs de transcrire les éléments phonétiques.
Un problème se pose alors : quelle orthographe employer ? L'orthographe phonétique est la meilleure, elle minimisera les erreurs de correction, mais les auditeurs utilisés ne la connaissent pas obligatoirement : d'où la possibilité d'erreurs de transcription. Pour éviter ce type d'erreurs, certains auteurs ont mis au point une méthode de "libre choix" [3.5.]
Lorsqu'un élément phonétique est émis, l'auditeur se voit proposer plusieurs choix possibles, c'est-à-dire la réponse exacte, plusieurs réponses erronées où l'on propose des logatomes phonétiquement proches du logatome émis, et une "solution nulle" correspondant à l'impossibilité de reconnaître le logatome émis. L'auditeur appuie alors sur un bouton dont le numéro correspond à la réponse choisie.
L'expérience a montré qu'un ensemble de 7 choix possibles donne des résultats analogues aux mesures par transcription, tout en éliminant des erreurs indépendantes des conditions acoustiques. Il faut noter que dans le même ordre d'idées, on a proposé d'utiliser le temps de réaction des auditeurs pour donner des mesures plus précises de l'intelligibilité, dans le cas de coefficients élevés [3.5.]

Nous ne considérerons que les mesures d'intelligibilité faites sur l'ensemble des logatomes, et non celles relatives à tel ou tel type de phonèmes, qui ont amené certains auteurs comme Pollack à définir des "matrices de confusion". [3.6.]

Revenons aux mesures classiques d'intelligibilité, où les auditeurs doivent transcrire les logatomes. Une étude de Stuckey [3.7.] a montré qu'un groupe d'auditeurs, appelé à écouter des listes de logatomes dans une certaine condition acoustique, et à transcrire les logatomes perçus sur une feuille de papier, améliore ses résultats au cours de séances successives (Fig. 3.2.)
En effet, les auditeurs s'habituent aux auditions de logatomes, s'habituent à leur transcription, et ont tendance à mémoriser, tout au moins partiellement, les logatomes.

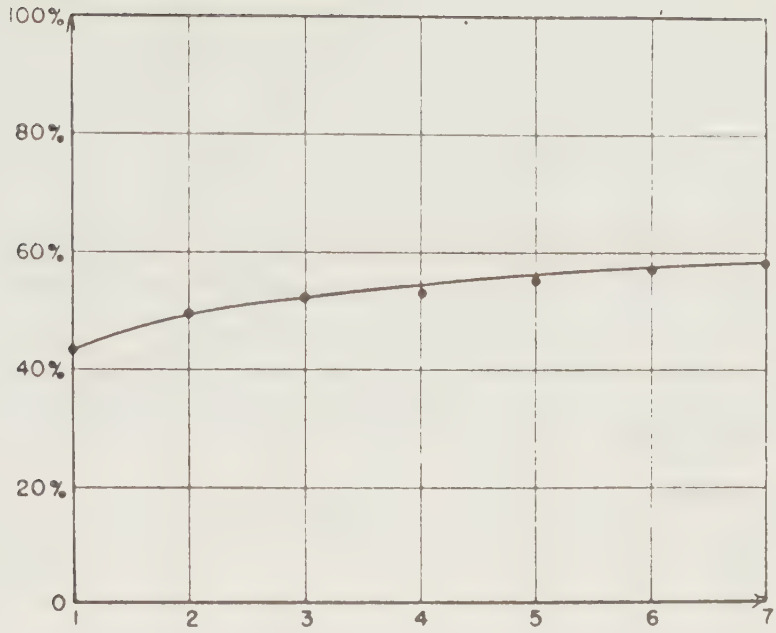


Fig. 3.2.- Influence du nombre de séances de mesure sur l'intelligibilité donnée par un groupe d'auditeurs, d'après Stuckey [3.7.]

Pour obtenir des résultats fiables, deux méthodes semblent possibles : si l'on peut renouveler souvent les auditeurs, et si l'on veut avoir toujours affaire à des auditeurs "naïfs", on s'expose à une certaine dispersion dans les mesures, à moins que le nombre d'auditeurs soit très élevé. D'autre part, les erreurs de transcription risquent d'être considérables. Dans ce cas, comme le faisait remarquer F. Canac [3.8.] , il est préférable d'utiliser des mots du langage courant, qui ne risquent pas de choquer l'auditeur, contrairement aux logatomes. Le nombre de listes doit être suffisant pour effectuer une mesure, ce qui dépend du nombre d'auditeurs disponibles. Par exemple, si l'on veut effectuer une mesure sur 1 000 logatomes écoutés, et si l'on dispose de 5 auditeurs nouveaux à chaque mesure, il suffira de constituer un petit stock de 200 logatomes. Evidemment, le problème qui consiste à trouver, ne serait-ce que 5 auditeurs nouveaux à chaque mesure, est assez difficile.

Une autre méthode, en pratique plus avantageuse, est d'utiliser des auditeurs entraînés; ceci suppose que ces auditeurs aient accompli un certain nombre de mesures préalables qui les amènent à donner, dans les mêmes conditions acoustiques, un coefficient d'intelligibilité maximal. Cependant, l'effet de mémorisation doit être limité, autant que possible. Il apparaît nettement, que si l'on utilise des phrases, le nombre de listes devra être important, pour que le matériau puisse être renouvelé assez souvent lors des mesures successives : on pourra constituer par exemple un ensemble de 100 listes de 50 phrases différentes. Si l'on utilise des mots, le stock de listes à constituer pourra être plus faible. Enfin, si l'on utilise des logatomes, leur nombre pourra être considérablement plus réduit.

Voici quelques exemples des différents paramètres utilisés par quelques auteurs pour des mesures d'intelligibilité [3.9] [3.10] [3.7.] [3.11.]

Nom	Année	Type de perturbation	nombre d'auditeurs	nombre d'éléments phonétiques
CHAVASSE	1950	bruit	5	1500
LEHMANN	1962	bruit	5	250
STUCKEY	1963	bruit	6	1800
JANUSKA	1968	réverbération	50	1000

Les éléments phonétiques utilisés étaient des logatomes, sauf dans le cas de Lehmann, qui a utilisé des mots et des phrases. On remarquera que, dans le cas de mesures d'intelligibilité dans une salle, on ne peut faire une mesure à une place donnée qu'avec un seul auditeur; quand on envisage plusieurs places, on doit avoir recours à des permutations circulaires d'auditeurs. Au contraire, dans le cas de mesures en présence de bruit, à l'aide d'écouteurs, il est possible d'effectuer une mesure simultanément sur plusieurs auditeurs, à condition que les écouteurs utilisés éventuellement soient identiques. Comme on peut le voir, les divers auteurs sus-nommés utilisent un nombre d'auditeurs et d'éléments phonétiques très

variables. Certains auteurs prennent soin de vérifier que les auditeurs ont une audition normale, et ne présentent aucun trouble auditif.

La "présentation" des mots ou des logatomes a, elle aussi, son importance. En effet, on peut émettre ces éléments phonétiques sans phrases introductives ou conclusives, et en veillant à les espacer assez pour qu'une transcription aisée soit possible. Cependant, de nombreux auteurs ont préféré utiliser un texte introductif et conclusif, méthode dont les avantages sont signalés par Kryter [3.12] :

"L'intérêt de la phrase porteuse est quintuple :

- Elle donne un moyen de contrôler le niveau de l'effort vocal.
- Elle assure un niveau normal et naturel au mot-test.
- Elle alerte l'auditeur.
- Elle procure un effet de masque par réverbération lorsque le test a lieu dans une salle.

La phrase porteuse doit être prononcée en moins de 2 secondes".

Le défaut irrémédiable d'une mesure d'intelligibilité est qu'elle dépend du type d'éléments phonétiques utilisés lors des différentes mesures, ceci conditionne les effets de mémorisation, cela la précision de la mesure. Des essais de normalisation ont été proposés [3.12] et il faut souhaiter que les auteurs s'y conforment, ou en tout cas indiquent clairement les divers paramètres de leurs mesures.

3.1.3.- Les locuteurs-

Les locuteurs sont chargés de prononcer les phrases, les mots ou les logatomes : on prend des sujets ayant une voix "normale", sans accent local particulier. On peut trouver judicieux de prendre des "speakers professionnels", par exemple. En fait, on peut se demander s'il n'est pas préférable d'utiliser un "locuteur moyen" dont la diction est bonne, sans être pour autant exceptionnelle. Pour étudier l'intelligibilité dans un amphithéâtre de faculté, la meilleure solution serait de faire lire les listes à des professeurs. On peut également utiliser des locuteurs masculins ou féminins. Cependant, s'il est rare que dans telle salle, une femme prenne la parole, on peut douter de l'intérêt d'utiliser des locuteurs féminins...

Malgré ces questions qui peuvent être soulevées, le problème est beaucoup plus simple que celui relatif aux auditeurs : un locuteur con-

venable n'est pas trop difficile à trouver : il faut également noter qu'il est souvent avantageux d'enregistrer une fois pour toutes l'ensemble des listes en chambre sourde, et de les rediffuser ici ou là, lors des mesures d'intelligibilité. Bien entendu, il faut choisir un diffuseur dont le rayonnement soit connu, sinon analogue à celui d'un locuteur réel.

3.1.4.- Méthode choisie pour nos mesures.-

Lorsque j'ai eu à effectuer des mesures d'intelligibilité, je me suis heurté au problème des auditeurs. On m'excusera d'employer ici la première personne du singulier, mais on comprendra qu'il m'était difficile de l'éviter. Il m'est apparu comme pratiquement impossible de "dresser" une équipe d'auditeurs : les enseignants ont la possibilité d'avoir recours à des étudiants de bonne volonté, ou heureux d'être agréables à leur professeur, mais tel n'était pas mon cas. Je n'avais pas de possibilité financière de payer quelques auditeurs extérieurs à l'institut. Il m'était difficile de débaucher plusieurs de mes collègues en mettant à contribution leur bonne volonté et leur patience. Comme j'était bien décidé à utiliser des auditeurs expérimentés, j'ai dû me décider à utiliser le seul auditeur taillable et corvéable à merci, et de plus, convaincu de l'intérêt de ce travail : moi-même.

J'ai composé, à partir des listes de logatomes français indiqués par Matras [3.13] 20 listes de 50 logatomes du type consonne-voyelle-consonne, soit 1000 logatomes au total. Ces logatomes se trouvent insérés dans de courtes phrases, ainsi le logatome "pat" dans la phrase : "écoutez pat maintenant". Le nombre total de logatomes a été estimé suffisant pour éviter des effets de mémorisation accentués. J'ai exclu des listes de logatomes les sons nasalisés tels que an, in, on, un, afin qu'il n'y ait pas d'équivoque sur la structure des logatomes envisagés. La correction a été faite en considérant un logatome comme faux, si l'un au moins des phonèmes était mal perçu. L'auditeur et le correcteur étant une même et unique personne, on peut estimer que les erreurs de correction sont très faibles. De même, les erreurs de transcription sont très limitées, l'auditeur connaissant bien la structure et le type de logatome émis. Pour certaines mesures, l'intelligibilité a été évaluée sur l'ensemble des 1000 logatomes (cas de la salle de conférences). Pour d'autres (chambre sourde), on s'est contenté de 200 logatomes, soit 4 listes prises au hasard parmi les 20.

3.2.- Méthodes de prévision de l'intelligibilité de la parole dans les salles.-

Dans ce chapitre, nous nous proposons de rappeler les différentes techniques élaborées jusqu'à ce jour pour estimer l'intelligibilité dans les salles, hormis les techniques basées sur les échogrammes, dont nous parlerons après. Le problème de la "prévision de l'intelligibilité" est la détermination, à l'aide des différents facteurs caractérisant les conditions acoustiques envisagées, d'un coefficient numériquement égal à l'intelligibilité mesurée à l'aide d'éléments phonétiques.

3.2.1.- Méthode de Knudsen [3.14]

Cette méthode est la première, et la plus ancienne à notre connaissance. Elle repose sur la constatation suivante : dans une salle, l'intelligibilité est d'autant plus faible que la durée de réverbération est plus longue, et que le niveau de la source est plus faible, en deçà d'une certaine limite. On sait que la notion de "durée de réverbération" dûe à Sabine, a été longtemps considérée comme la caractéristique essentielle de l'acoustique d'une salle. Il était donc logique que Knudsen se soit basé sur ce paramètre pour ses estimations de l'intelligibilité.

D'après cet auteur, l'intelligibilité It serait un produit de facteurs indépendants : la formule proposée est la suivante :

$$It = 0{,}96.k_E .k_R .k_B .k_F$$

le facteur k_E est relatif au niveau du signal émis, k_R à la durée de réverbération, k_B au niveau du bruit, et k_F un facteur de correction, qui, suivant Knudsen, dépend de la forme de la salle. Tous ces facteurs sont, bien entendu compris entre 0 et 1. Le facteur 0,96 tient compte du fait que, lors des mesures de Knudsen, les auditeurs n'ont jamais donné un pourcentage de logatomes correctement transcrits supérieur à 96. Bien entendu, ce facteur devrait être égal à 1, ou très légèrement inférieur, pour des auditeurs rompus aux mesures d'intelligibilité. Les différents facteurs ont été déterminés expérimentalement, dans plusieurs salles dont Knudsen fai-

sait varier les propriétés absorbantes en disposant plus ou moins de laine de verre. Le facteur k_E est égal à 1 si le niveau émis est de 70 dB. Le facteur k_R décroît, à partir de 1, si la durée de réverbération dépasse la valeur considérée comme optimale, en fonction du volume de la salle considérée. Le facteur k_B dépend du bruit ambiant. Enfin, le facteur k_F peut paraître assez mystérieux : il sert à caractériser, par opposition à une salle "normale" rectangulaire, une salle dont l'acoustique serait plus défavorable, en présentant par exemple des effets de focalisation particulièrement fâcheux. A ce sujet, on peut citer les salles circulaires, ou même les salles en éventail.

Comme on le voit, la méthode de Knudsen ne tient guère compte de l'acoustique propre d'une salle qu'à l'aide de la notion de "durée de réverbération", corrigée le cas échéant par un "facteur de forme" assez vague. Malgrè ces réserves, la méthode de Knudsen a l'avantage d'avoir été la première proposée pour appréhender une "prévision" de l'intelligibilité.

3.2.2.- Méthode des indices d'articulation.-

Les téléphonistes ont défini un coefficient analogue au coefficient d'intelligibilité, qui a le même domaine de variations, et dont l'intérêt est d'être un indice cumulatif. En effet, si l'on considère l'influence d'un filtrage passe-haut, et celle d'un filtrage passe-bas sur l'intelligibilité, il apparaît que l'intersection des deux courbes a lieu pour une fréquence de 2000 Hz, et pour une valeur de l'intelligibilité égale à 0,68. Or, si chaque bande de fréquence 100-2000 Hz et 2000-10000 Hz donne une intelligibilité de 68%, les deux bandes réunies donnent une intelligibilité de 100%. On voit donc que le coefficient d'intelligibilité n'est pas un indice sommable. Au contraire, si l'on adopte un coefficient dont la valeur est de 0,5 lorsque l'intelligibilité est de 68%, on peut définir ainsi un coefficient sommable, l'indice d'articulation. D'autres valeurs de cet indice sont obtenues en utilisant des listes de logatomes mélangées à du bruit, et dont l'intelligibilité est déjà plus faible, et en leur faisant subir un filtrage analogue, passe-haut et passe-bas (Fig. 3.3 et 3.4)

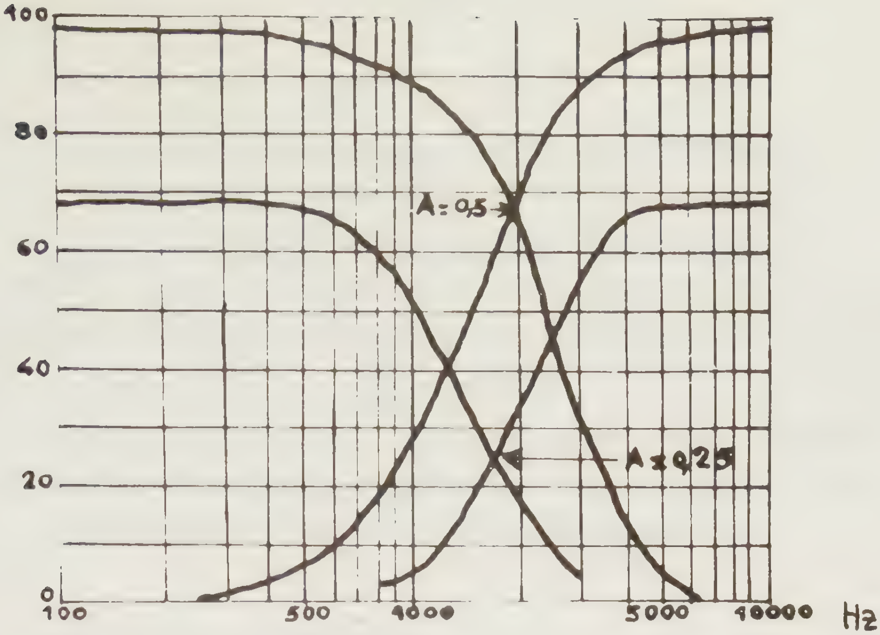
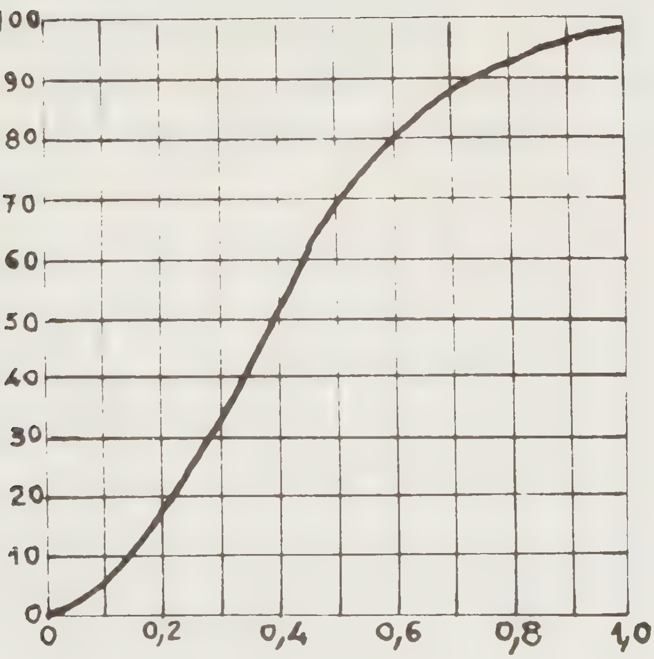


Fig 3.3. Définition de l'indice d'articulation A à partir de l'intelligibilité de la parole filtrée, d'après French et Steinberg [3.15]



[illegible] d'articula-tion [illegible] French et [illegible] [3.15]

Dans le cas des télécommunications, l'indice d'articulation a été très utilisé, en découpant la bande utile du spectre de la parole en plusieurs régions d'égale contribution à l'intelligibilité : French et Steinberg ont proposé une division en 20 bandes de fréquences, dont chacune apporte une contribution $a_i \leqslant 0{,}05$ suivant le masquage de chaque région par la bande correspondante du bruit perturbateur [3.15] . L'indice d'articulation est alors $a = \sum_i a_i$. Certains auteurs ont préconisé l'utilisation de l'indice d'articulation, ou d'un indice cumulatif, dans le cas où la parole est masquée par des effets de réverbération.

a) Méthode de Bolt et Mac Donald. [3.16]

Cette méthode, proposée en 1947, suppose que la parole peut être représentée par une succession d'impulsions sonores, dont la durée est 0,13s, alors que celle des intervalles les séparant est de 0,1s. L'intensité de chaque impulsion peut être comprise entre 0 et 30 dB (Fig. 3.5).
Ces données découlent de mesures expérimentales. Chaque impulsion peut être masquée par les impulsions précédentes. On découpe l'impulsion considérée en tranches horizontales, suivant les possibilités de masquage dûes aux impulsions précédentes : ce masquage est déterminé d'après la durée de réverbération. Chaque tranche apporte à l'indice d'articulation une contribution $a_i\, p_i$, où a_i est la fraction de hauteur de cette tranche ($a_i = h_i/30$), et p_i sa probabilité d'échapper au masquage.
Ainsi, on a :

$$a_1 = h_1/30 , \qquad p_1 = 1,$$

ce qui signifie que le sommet de l'impulsion échappe au masquage : on définit un indice total $A = \sum a_i p_i$, dont les variations sont comparables à celles de l'indice d'articulation déduit des mesures d'intelligibilité.

En 1950, Korn a repris cette estimation, qui était d'autant moins satisfaisante que la durée de réverbération était plus élevée, en introduisant un nouvel élément, le rapport de l'énergie du son direct à celle du son réverbéré, ce qui revient en fait à modifier la forme des impulsions schématiques [3.17] . Avec cette modification, Korn a trouvé des valeurs dont l'accord avec les expériences de Knudsen est nettement plus satisfaisant. Cependant, l'ensemble des résultats ne s'appuie pas sur des expériences nombreuses, et apparemment, ce travail n'a pas eu de suites.

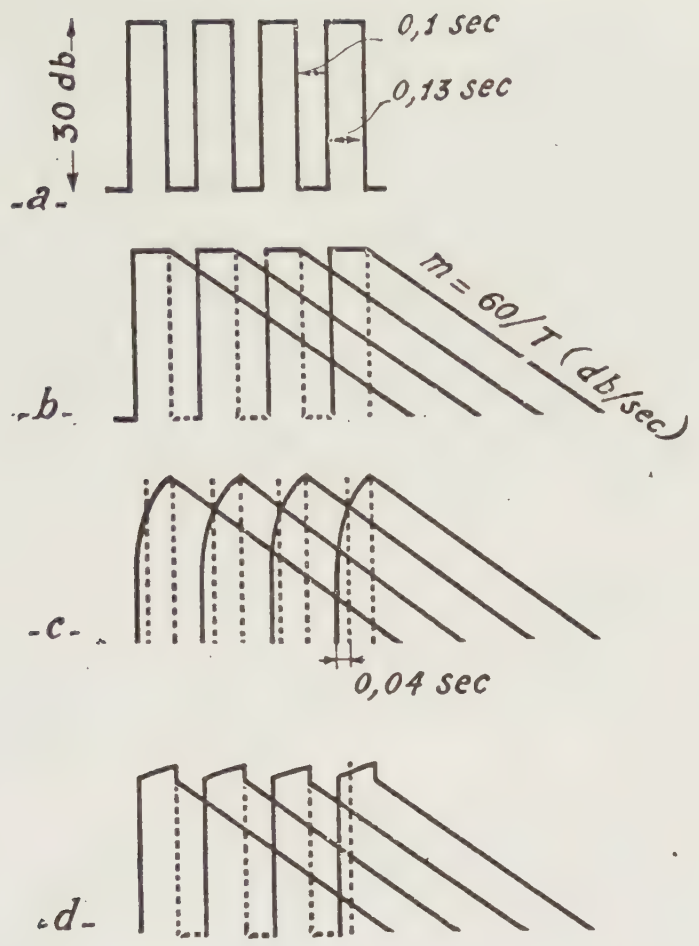


Fig 3.5. Représentation de la parole sous la forme d'impulsions élémentaires moyennes, selon Korn et Bolt [3.16] [3.17]

a) allure des impulsions à la source

b) déformation à la disparition par la réverbération

c) déformation à l'attaque et à la disparition

d) influence de l'action du son direct

b) Méthode de Janssens.

L'indice calculé par Bolt, ou par Korn, n'avait avec l'indice d'articulation de French et Steinberg que la propriété commune d'être cumulatif : mais son calcul en était fort éloigné, puisque cet indice était relatif à un découpage des niveaux, alors que l'indice d'articulation relève d'un découpage des fréquences. Le travail de Janssens (1957) est justement relatif à l'utilisation du véritable indice d'articulation, adapté au cas de la réverbération [3.18] . La bande de fréquences est divisée en 5 bandes seulement, dont chacune donne une contribution $a_i\ r_i$. Le facteur a_i est déduit, pour chaque bande, de la différence S_i entre la valeur du son direct et celle du son réverbéré. Le facteur r_i dépend également de S_i et de la durée de réverbération.

Les valeurs de a_i et r_i sont déduites d'abaques construits à l'aide de résultats expérimentaux. Les résultats obtenus avec ces prévisions montrent une concordance assez satisfaisante avec les mesures d'intelligibilité, à 10% près.

En résumé, avec la méthode de Korn comme avec celle de Janssens, on fait figurer deux paramètres d'une salle : la durée de réverbération, et le rapport du son direct au son réverbéré. Cependant, le niveau intrinsèque du son direct n'y figure pas, ce qui laisse penser qu'aucune de ces méthodes ne peut donner une prévision précise. Nous aurons par la suite l'occasion de voir quel rôle déterminant le niveau du son utile peut avoir dans les problèmes d'intelligibilité.

3.2.3.- Les procédés de corrélation.-

Comprendre un mot, c'est le reconnaître : nous avons déjà eu l'occasion de parler de ce processus de reconnaissance : on conçoit bien quel est le processus post-auditif qui entre ne jeu dans l'opération d'intelligibilité; il s'agit bien d'une comparaison d'un élément phonétique, long ou bref, avec une multitude d'éléments préalablement enregistrés dans la mémoire. L'auditeur de logatomes, s'il est averti, connaît la structure de l'élément phonétique qu'il va recevoir : il y a également un phénomène de reconnaissance, même à un niveau moins élevé, alors que ce phénomène devient extrêmement hasardeux, quant à son résultat, si l'on en-

visage des éléments phonétiques indéterminés.
Il existe une opération mathématique bien connue qui permet de comparer deux fonctions $x(t)$ et $y(t)$, c'est l'opération de corrélation.
En fait, cette opération semble d'autant mieux schématiser les processus cérébraux qu'on peut estimer que l'auditeur percevant un signal acoustique $y(t)$ le comparera à plusieurs éléments préenregistrés $X(t)$ et choisira celui que ressemble le plus à $y(t)$, soit $x(t)$.

Supposons qu'on ait échantillonné chaque fonction $x(t)$ et $y(t)$, représentant des pressions acoustiques, en n valeurs x_i et y_i [3.19] .
On considère l'écart quadratique moyen :

$$z(a) = (1/n) \sum_i (x_i - a y_i)^2$$

où a est un coefficient tel que $z(a)$ soit minimum, soit $z'(a) = 0$:

$$a = \frac{\frac{1}{n} \sum_i x_i \, y_i}{\frac{1}{n} \sum_i y_i^2}$$

Dans ces conditions $z_{min} = (1/n) \sum_i x_i^2 \, (1 - \rho^2)$ où

$$\rho = \frac{(1/n) \sum_i x_i \, y_i}{\sqrt{(1/n) \sum x_i^2 \; (1/n) \; y_1^2}}$$

Z_{min} sera nul si $\rho = 1$, dans ce cas les deux fonctions $x(t)$ et $y(t)$ seront totalement corrélées. Au contraire, si ρ est nul, z_{min} est maximal, dans ce cas les deux fonctions $x(t)$ et $y(t)$ ne sont pas corrélées.
ρ est le coefficient de corrélation normé, qui en utilisant directement les fonctions $x(t)$ et $y(t)$ aurait pour forme :

$$\rho = \frac{\overline{x(t).y(t)}}{\sqrt{\overline{x^2}(t).\overline{y^2}(t)}} \qquad \text{avec} \qquad \overline{x}(t) = \frac{1}{T}\int_0^T x(t)\ dt$$

$$\overline{x^2}(t) = \frac{1}{T}\int_0^T x^2(t)\ dt$$

et des relations analogues pour $y(t)$, où T est la période d'observation des fonctions $x(t)$ et $y(t)$.

Au lieu d'effectuer l'opération $(1/T)\int_0^T x(t).y(t)\ dt$, il peut être intéressant d'effectuer un décalage τ de la fonction $y(t)$ dans le temps. On obtient alors une fonction de corrélation :

$$C_{xg}\ (\tau) = (1/T)\int_0^T x(t).y(t-\tau)\ dt$$

si $y(t) = x(t)$, on obtient une fonction d'autocorrélation :

$$C_{xx}\ (\tau) = (1/T)\int_0^T x(t).x(t-\tau)\ dt$$

Cette fonction est importante, car il est fort probable que le cerveau ait recours à une opération analogue pour extraire l'information utile du bruit, ou du brouillage qui lui est mélangé, ce qui est une propriété très intéressante de l'opération d'autocorrélation.

Fourdouev a mesuré des fonctions d'autocorrélation pour différents types de signaux, et il en a déduit des courbes de cohérence. [3.20] En ce qui concerne la parole, il est assez caractéristique que la chute de la fonction est assez brutale. Au contraire, pour d'autres types de signaux, la décroissance est beaucoup plus lente (Fig. 3.6).
A titre indicatif, nous avons effectué quelques mesures de corrélation et d'autocorrélation sur des signaux réverbérés et non réverbérés. Au préalable, on avait enregistré dans une salle de conférences sur laquelle nous reviendrons plus loin, une liste de logatomes sans phrases introductives ni conclusives. Un premier micro placé devant le locuteur enregistrait sur la première voie d'un magnétophone bipiste le signal considéré comme pur,

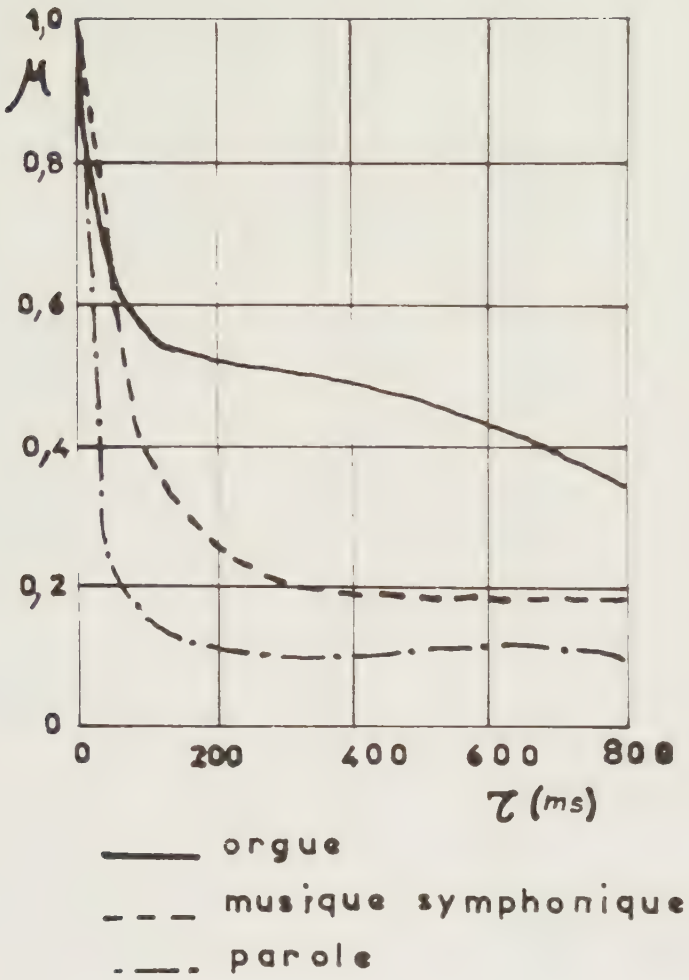


Fig. 3.6 - Courbes de "cohérence" d'un signal avec lui-même, retardé d'un délai , d'après Fourdouev [3.20]

soit x(t) . Un second micro, placé dans la salle à 5m environ du locuteur enregistrait sur la deuxième voie le signal réverbéré y(t).

Nous avons tracé une fonction d'autocorrélation C_{xx} à l'aide d'un corrélateur SAIP, pour quelques logatomes : les figures 3.7 et 3.8 donnent la courbe obtenue pour le logatome "pob".On obtient une courbe pseudo-périodique, la parole étant elle-même de nature pseudo-périodique [3.21] .

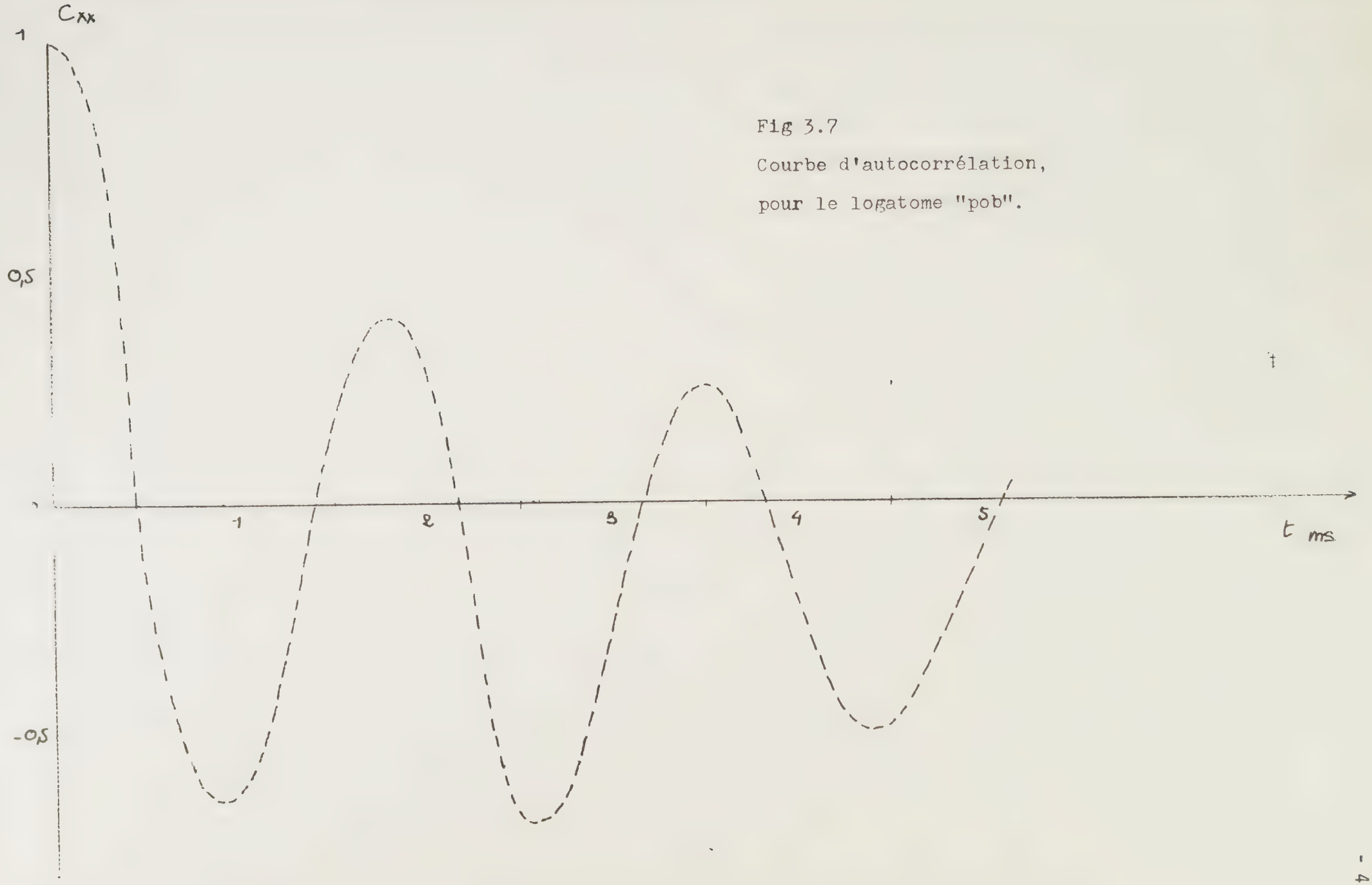


Fig 3.7

Courbe d'autocorrélation,

pour le logatome "pob".

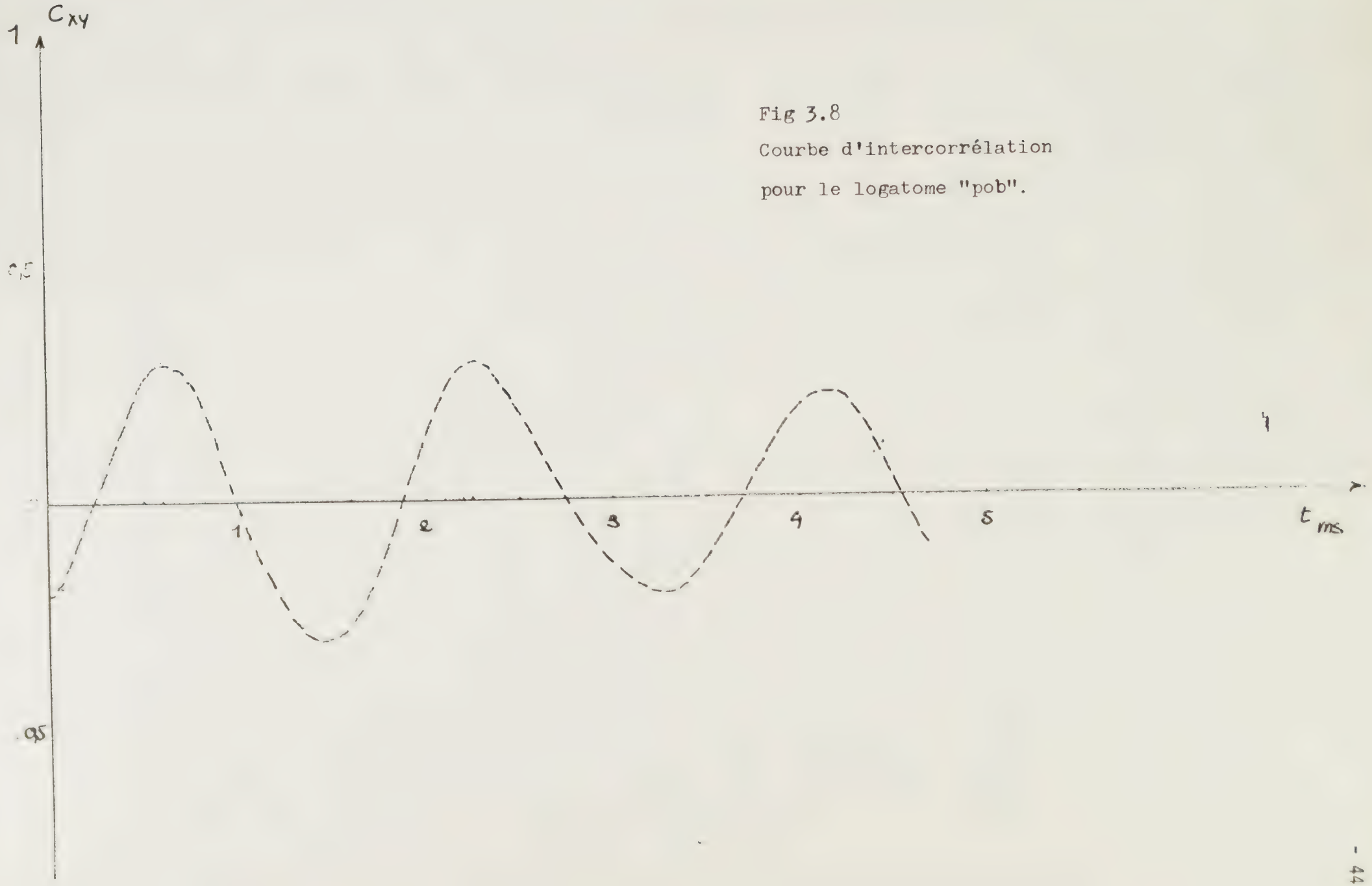


Fig 3.8
Courbe d'intercorrélation
pour le logatome "pob".

L'amplitude maximale des oscillations, égale à 1 pour $\tau = 0$, tombe à 0,67 si τ est compris entre 10 et 20 ms, et à 0,4 si τ est compris entre 20 et 30 ms. La fonction d'autocorrélation C_{yy} a une allure très proche de C_{xx} . Mais la fonction d'intercorrélation C_{xy} a une amplitude qui ne dépasse pas 0,3 , tout en ayant une pseudo-période comparable à celle de C_{xx} ou de C_{xy}.
Si τ est supérieur à 50 ms, l'amplitude de C_{xy} diminue nettement.

La pseudo-période dépend du logatome considéré, et semble-t-il, de la première région formantique de la voyelle : ainsi, pour le logatome "pob", on a une pseudo-fréquence de 580 Hz.

Ces résultats se retrouvent, grosso-modo, pour tous les logatomes, avec des pseudo-fréquences légérement différentes : pour "bour" , par exemple, la pseudo-fréquence est de 526 Hz. Il apparaît bien que la corrélation se détériore assez rapidement, et que l'intercorrélation est assez faible, alors que le signal réverbéré $y(t)$ est intelligible à 90%.

Nous avions fait, il y a quelques années, une autre expérience du même type, dont les résultats peuvent paraître plus satisfaisants. Nous avions fait subir au préalable un "codage" aux signaux $x(t)$ et $y(t)$ [3.21] . De nombreux auteurs, dont les travaux sont mentionnés dans notre diplôme, ont montré qu'il est possible de transformer la parole en une succession de signaux rectangulaires de même amplitude, et de durée variable, et dont les discontinuités correspondent aux passages à 0 du signal p(t) préalablement préaccentué, c'est-à-dire filtré dans la zone formantique. L'intelligibilité de la parole ainsi transformée reste élevée, de l'ordre de 90% pour les logatomes. En mesurant le nombre de passages à zéro du signal $p(t)$ par unité de temps $\Delta\tau$, et en prenant par exemple $\Delta\tau$ = 20 ms, on obtient une distribution $x_i(t)$ pour le signal original, soit $x_1(t)$ pour $t \in (0, \Delta\tau)$, $x_2(t)$ pour $t \in (\Delta\tau, 2\Delta\tau)$ etc... De la même façon, on obtient une distribution $y_i(t)$ pour le signal réverbéré. On peut déterminer un coefficient de corrélation entre les deux distributions $x_i(t)$ et $y_i(t)$.
Les résultats mettent en lumière un fait important : la cohérence des deux signaux diminue, même si le signal $y(t)$ n'a subi que de petites altérations par rapport à $x(t)$, dues par exemple aux seuls appareils de mesure (Fig. 3.9).La situation acoustique repérée (S) correspond à une copie de la bande originale, effectuée dans une chambre sourde, celle repérée (SC) est

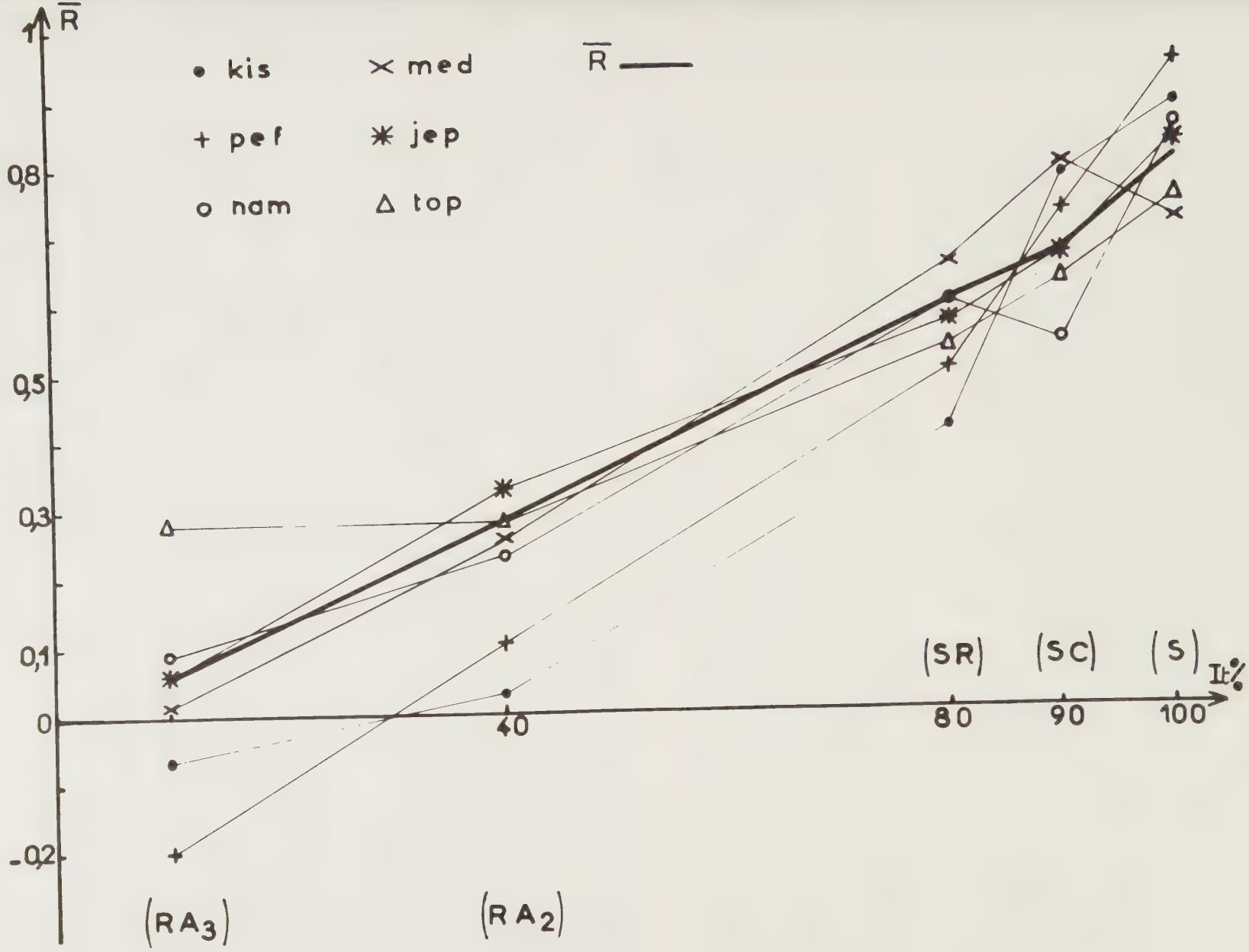


Fig 3.9. Variation du coefficient de corrélation pour divers locuteurs en fonction de l'intelligibilité.

relative à un enregistrement dans la salle de conférences, (SR) désigne une copie effectuée dans une salle réverbérante, (RA2) et (RA3) des enregistrements effectués avec un procédé de réverbération artificielle. Les niveaux globaux des différents signaux sont identiques.

Il apparaît bien que la corrélation varie dans le même sens que l'intelligibilité, cependant, si l'intelligibilité varie de 10% à 100%, la corrélation est située entre 0,07 et 0,75 , par conséquent la variation est plus réduite. D'autre part, il apparaît une certaine dispersion dans les résultats, dont on peut se demander si elle est significative ou non.

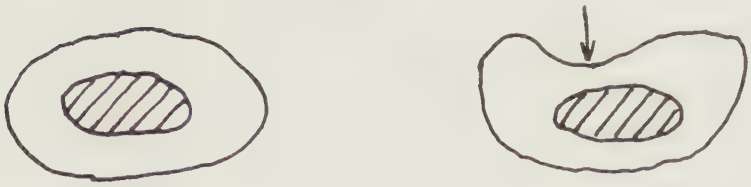
Par conséquent, si une technique de corrélation peut donner des résultats globaux sur la déformation d'un signal, et si cette déformation a une certaine relation avec l'intelligibilité, nous ne sommes pas convaincus d'avoir là un moyen d'analyse particulièrement précis pour le domaine très particulier de l'intelligibilité. Peut-être la corrélation est-elle mieux adaptée pour étudier la "qualité" d'une salle pour la musique.

Pour revenir à la notion schématique d'information sémantique et d'information esthétique, nous dirons qu'une diminution de la cohérence est relative à une modification de l'information véhiculée par le message : cependant, si par exemple sa coloration, liée à la partie esthétique de l'information peut avoir changé, l'information sémantique, à laquelle est liée l'intelligibilité, peut avoir conservé toute son intégrité : la diminution de la cohérence risque de chiffrer d'abord une altération de l'information esthétique.

Les dessins naïfs de la figure 3.10 schématisent le problème : l'information totale est représentée par une forme dont le noyau est constitué par l'information sémantique. Une certaine déformation modifie cette forme et altère l'information esthétique sans modifier l'information sémantique : l'intelligibilité reste de 100%, le coefficient de corrélation peut être réduit à 0,7. Envisageons cette fois une perturbation beaucoup plus importante : l'information esthétique est déformée, mais l'information sémantique est également atteinte : l'intelligibilité diminue, la cohérence devient encore plus faible.

Une autre méthode de corrélation qui donnerait des résultats in-

1) perturbation faible



2) perturbation importante

Altérations de l'information dues à des perturbations

1) seule l'information esthétique est déformée.

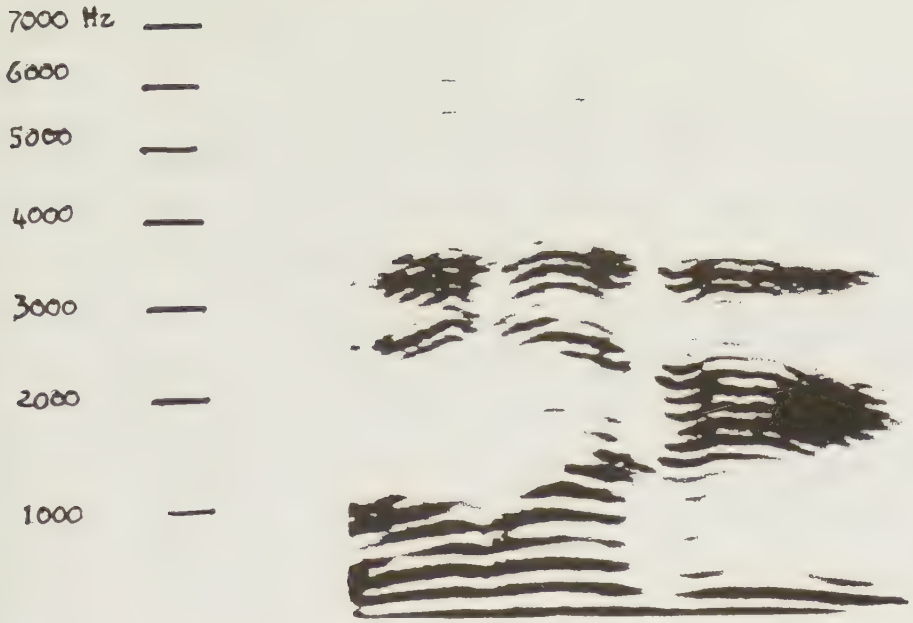
2) l'information sémantique est atteinte également.

Fig 3.10

téressants est celle proposée par Leipp [3.23] . Aux signaux x(t) et y(t) utilisés directement lors de nos manipulations, ou lors de celles de Fourdouev, on substitue des images sonagraphiques. Par un procédé optique à la fois simple et ingénieux, Leipp a montré comment on pouvait obtenir la partie commune des deux images, ainsi que l'originalité de l'une et de l'autre par rapport à cette partie commune. Appliquée au cas de la réverbération, cette méthode montre ce qu'une salle apporte au message x(t) et ce qu'elle lui enlève. Mais une fois de plus, nous craignons qu'en essayant d'appliquer cette technique au problème de l'intelligibilité, on n'obtienne qu'un résultat très partiel, ou pire, qu'on se fourvoie.
Nous donnerons à ce sujet l'exemple d'un enregistrement au sonagraphe du groupe phonétique "au rebus", prononcé en chambre sourde, puis recopié dans notre salle de conférences (SC) (Fig. 3.11).

De toute évidence, la forme sonagraphique est bien différente d'un enregistrement à l'autre. Cependant, la salle de conférences envisagée est une bonne salle, où l'intelligibilité est de l'ordre de 90%.
On peut donc penser que la partie commune aux deux sonagrammes contient l'information sémantique du message utilisé, et que cette information reste peu altérée par rapport à l'information sémantique originale.
Mais, ici encore, il faut rester prudent : informations esthétique et sémantique sont intimement liées sur le plan de la structure physique : à vouloir les séparer artificiellement, on risque de faire des contresens.
Le problème du brouillage de l'information sémantique est loin d'apparaître clairement sur un sonagramme, car les possibilités d'extraction d'un message dans un bruit opérés par l'oreille semblent très supérieures à celles de l'oeil pour un sonagramme réverbéré.

En conclusion, on peut dire que les techniques de corrélation ont l'intérêt de schématiser un mode opératoire qui intervient fort probablement au niveau du cerveau, et qui permet une reconnaissance des formes [3.24] .
Ce procédé intervient, d'une part, pour extraire l'information utile d'un bruit, ou d'une perturbation analogue (autocorrélation),et d'autre part, pour comparer le signal reçu y(t) à un message primitivement enregistré dans la mémoire (intercorrélation).



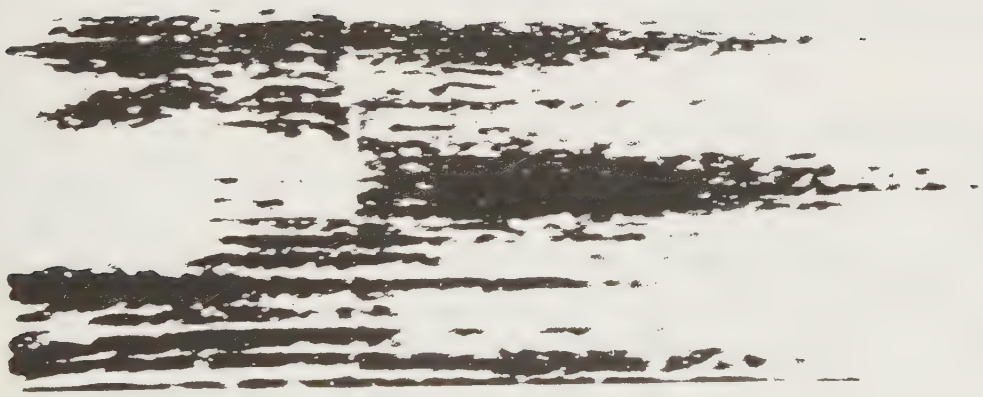


Fig 3.11. Groupe phonétique "au rebut" enregistré en chambre sourde (1) et en salle de conférences (2). Les échelles sont les mêmes pour les deux enregistrements. Durée de l'enregistrement original: 700 ms.

Si la comparaison optique des sonagrammes direct et réverbéré constitue un procédé digne d'intérêt, il ne s'agit certes pas d'une panacée, et nous ne pensons pas qu'il puisse résoudre d'une façon définitive le problème de l'intelligibilité de la parole.

Ajoutons encore qu'il ne faut pas perdre de vue le problème du niveau du signal reçu, qui est capital pour l'intelligibilité : or, les techniques de corrélation font abstraction de cet élément. Dans une salle, à un endroit donné, l'intelligibilité peut être mauvaise, non pas à cause d'une réverbération excessive, mais parce que le niveau sonore du signal reçu est trop faible.

3.2.4.- Méthode de Houtgast et Steeneken.-

Une méthode très originale a été récemment proposée par les néerlandais Houtgast et Steeneken [3.25] . Ces auteurs proposent, pour prévoir l'intelligibilité, d'envisager la déformation de l'enveloppe de la fonction $I(t)$, où I est l'intensité de la parole.
Ils ont montré que cette courbe a un spectre situé entre 0,25 et 25 Hz. La réverbération, ou le bruit, ont pour effet d'atténuer la modulation qui affecte la parole; si l'enveloppe du signal original a pour expression :

$i\ (1 + \cos 2\pi Ft)$, celle du signal réverbéré sera :

$i\ (1 + m \cos 2\pi F(t + \tau))$, où m est un coefficient compris entre 0 et 1, τ étant un retard quelconque. F est une fréquence comprise entre 0,25 et 25 Hz. La mesure est effectuée pour 21 valeurs de F réparties par bandes de tiers d'octave entre ces deux valeurs limites. On obtient alors 21 valeurs de MT, avec :

$$MT = 20 \log m$$

MT est appelé "fonction de transfert de modulation" (Fig. 3.12).

Il faut noter que la fréquence du signal modulé n'intervient pratiquement pas : des mesures ont été effectuées avec des bandes de bruit d'un octave centrées sur des fréquences de 500 Hz et 2000 Hz, sans donner de résultats différents.

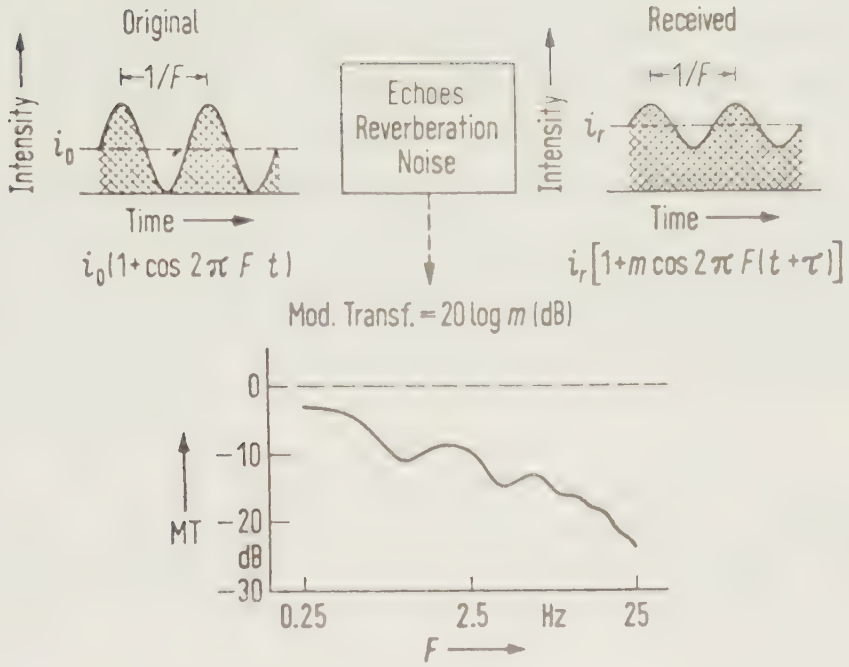


Fig. 3.12 - Définition de la "fonction de transfert de modulation", d'après Houtgast et Steeneken [3.25]

Une "fonction de transfert de modulation pondérée" est obtenue en sommant les contributions des différentes bandes de fréquences F suivant l'expression :

$$WMTF = \sum_{n=1}^{21} W_n \frac{1}{1 - k\,(MT)_n}$$

où W_n et k sont deux coefficients que l'on ajuste à l'aide des résultats expérimentaux. On a choisi la valeur $k = 0{,}3$ qui minimise l'écart entre la fonction WMTF et les mesures d'intelligibilité, effectuées sur des mots. D'autre part $W_n = 100/21$, ce qui donne $WMTF = 1$, pour $MT_n = 0$ dB , c'est-à-dire en l'absence de perturbations. L'écart entre les prévisions et les mesures est de l'ordre de 4,8%.

Les résultats avancés par les auteurs semblent montrer que cette méthode est efficace. Une question se pose : à quoi correspond, physiquement, cette "fonction de transfert de modulation"? Houtgast et Steeneken font remarquer qu'il s'agit en fait de la transformée de Fourier de la courbe de réverbération. Cependant, on peut reprocher à cette méthode d'être passablement sophistiquée, et quelque peu parachutée. D'autre part, on fait abstraction, une fois de plus, du niveau du son reçu, et on réduit le rôle de la salle à celui donné par la durée de réverbération.
En bref, cette méthode nous semble assez éloignée de la réalité physique; c'est pourquoi, sans nier son efficacité, nous ne lui accorderons pas une confiance illimitée.

Nous envisagerons maintenant des méthodes que nous appelerions volontiers "naturelles", ce sont celles basées sur les échogrammes, dont nous devons retracer l'historique, avant même d'étudier ses développements.

3.2.5. - Méthodes basées sur les échogrammes.-

a) Les échogrammes -

Le concept d'échogramme, ou diagramme d'échos, a été mis en lumière vers les années 1950, par les chercheurs de l'Université de Göttingen. On peut estimer qu'après les travaux de Sabine et la définition de la durée de réverbération, après les travaux complémentaires de Millington et Eyring, c'est l'acoustique ondulatoire qui s'était beaucoup développée avec Morse, Bolt, Maa, etc... Pour étudier la qualité d'une salle, les chercheurs de Göttingen se sont orientés à nouveau dans une optique qui relève nettement de l'acoustique géométrique.

La notion d'écho est celle qui s'impose de la façon la plus évidente au grand public : les profanes disent d'une salle qu'elle "a de l'écho", que "l'écho dure longtemps", etc... M. Jessel citait un jour le mot d'une célèbre comédienne, aujourd'hui disparue, et qui, parlant d'une salle de théâtre, disait : "L'acoustique y était si mauvaise qu'un écho arrivait avant même qu'on ait ouvert la bouche..."

La notion d'écho est associée au fait suivant : dans une salle, où l'on a défini la situation de la source et celle du récepteur, le son qui atteint le récepteur est non seulement le son direct, qui va de la source au récepteur, mais également le son réfléchi une ou plusieurs fois sur les différentes parois de la salle.

Nous verrons plus loin comment on peut, à l'aide des principes de l'acoustique géométrique, et en admettant une réflexion spéculaire du son, prévoir la répartition des différentes réflexions qui suivent le son direct, dans les coordonnées intensité-temps. Une question se pose alors : si l'on admet que la qualité d'une salle est déterminée en grande partie par ce diagramme de réflexions, comment peut-on obtenir expérimentalement un tel enregistrement? de toute évidence, c'est en émettant un son de durée très réduite : plus cette durée sera brève, plus on aura de chances de pouvoir distinguer les réflexions les unes des autres, tout au moins pour des délais relativement faibles.
Il est évident que si l'on émet un son permanent d'expression

$f(t) = h(t) \sin \omega t$ où $h(t)$ est la fonction de Heaviside, on obtiendra un établissement du régime permanent, qui risque de donner fort peu de renseignements sur la répartition des réflexions [3.26]
On peut émettre un train d'onde de forme :

$$f(t) = \left\{h(t) - h(t-\tau)\right\} \sin\omega t.$$

L'enregistrement obtenu dépendra d'une part de la durée τ , d'autre part de la fréquence ω utilisée.

En effet, les coefficients de réflexion des parois dépendent en général de la fréquence, et suivant le spectre de fréquence du train d'ondes émis, le diagramme de réflexion obtenu sera différent (Fig.3.13).
En fait, on recueille expérimentalement une pression sonore, et pour détecter dans les meilleures conditions les échos, on tâche d'utiliser un train d'onde très court, réduit par exemple à une période de sinusoïde à 1000 Hz.
Bien entendu, le choix de la durée de l'impulsion peut dépendre de la salle qu'on envisage d'étudier : si la salle est grande, les différentes réflexions seront relativement espacées et l'on peut se contenter d'une durée d'impulsion assez large; au contraire, si la salle est petite, on sera amené à choisir une durée d'impulsion assez brève pour obtenir une séparation suffisante des premières réflexions. Nous aurons l'occasion de revenir plus loin sur le problème de l'impulsion à émettre; pour l'instant, nous nous contenterons de dire que l'excitation est une impulsion acoustique brève, de durée inférieure à 5 ms. La figure 3.14 donne un exemple d'échogramme que nous

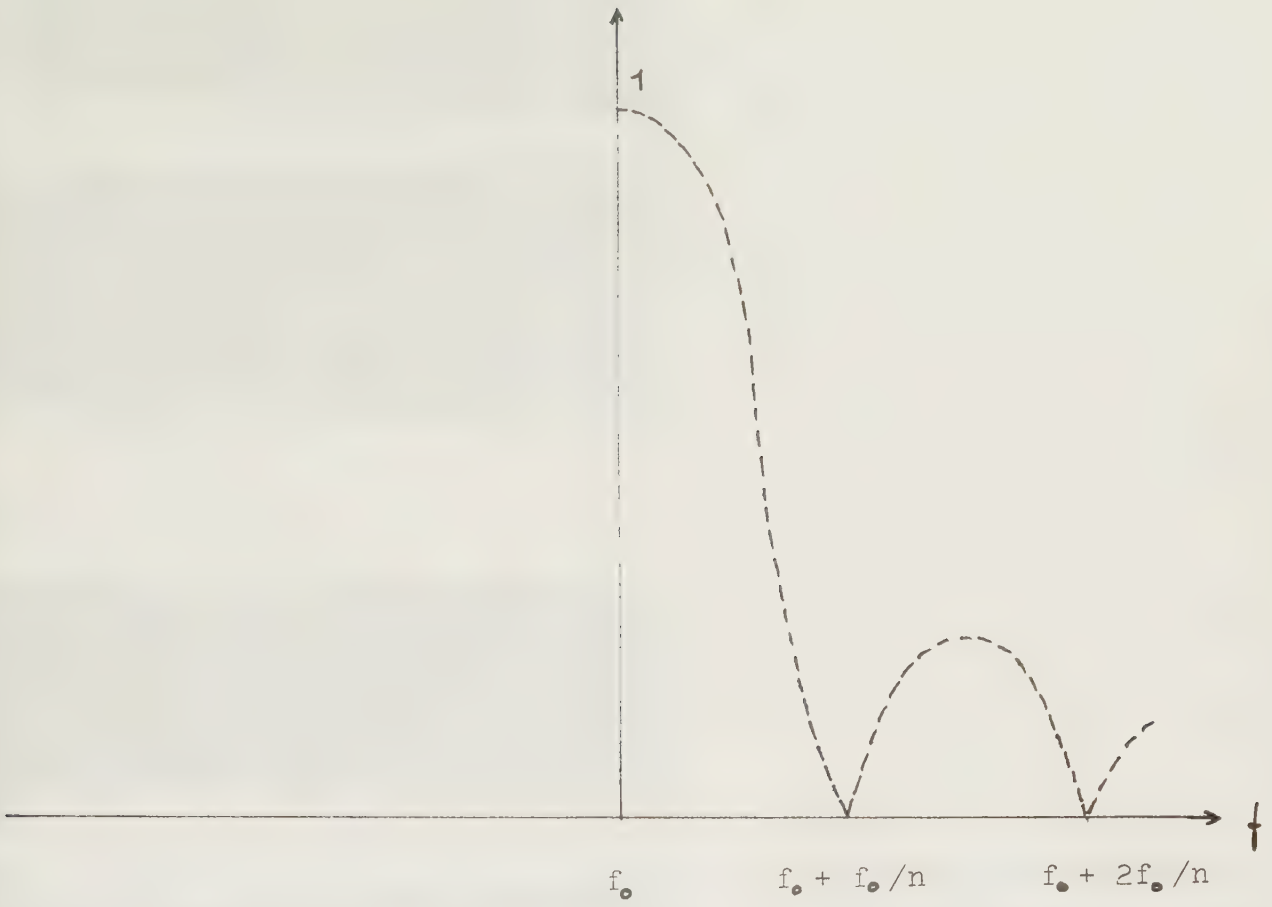
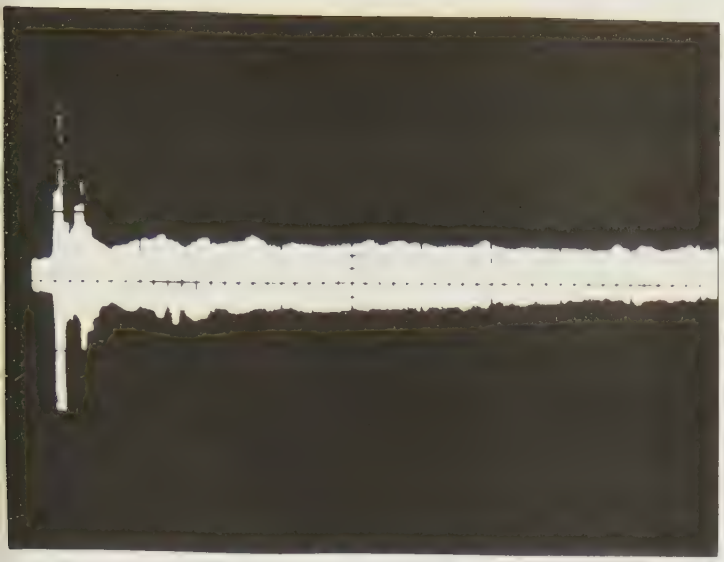
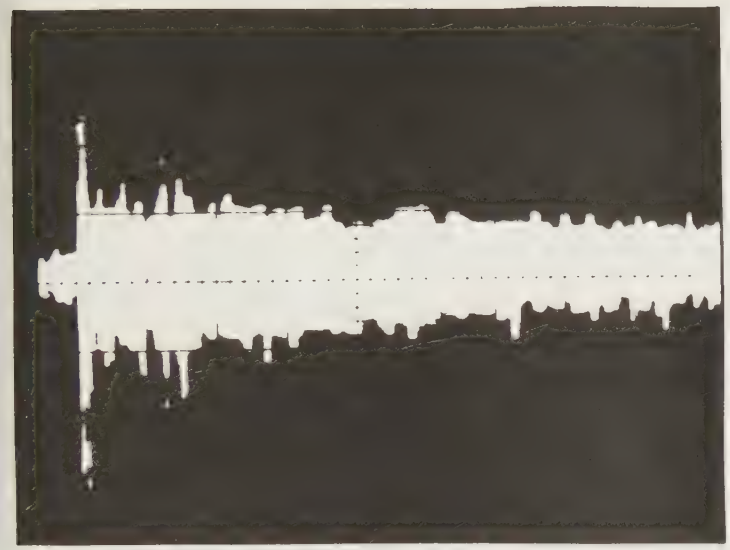


Fig 3.13. Spectre du signal $\{h(t) - h(t-\tau)\}$ sin $\omega_o t$
avec $\omega_o = 2\pi f_o$ et $\tau = nT_o$ (avec $T_o = 1/f_o$)
Le niveau des maximums est égal à $1/\pi(k+1/2)$, avec k entier positif. La courbe est symétrique par rapport à $f = f_o$.





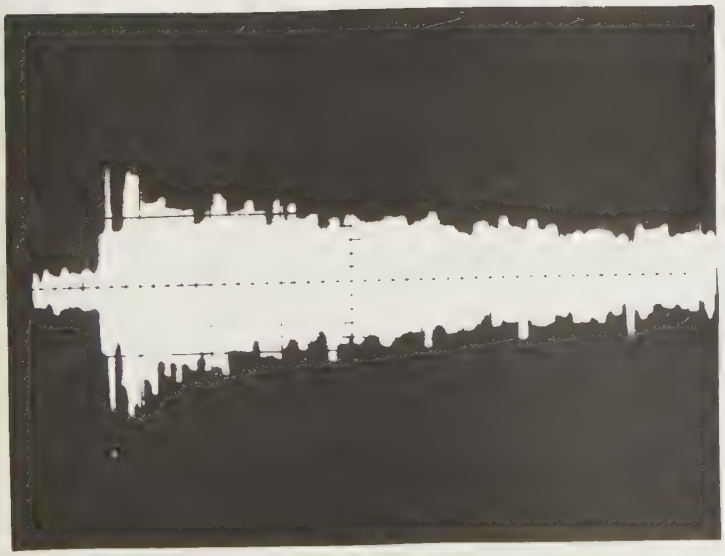


Fig 3.14
Exemples d'échogrammes enregistrés dans l'abbaye du Thoronet.
Echelle horizontale:
100 ms par carreau.

avons enregistré dans l'église de l'abbaye du Thoronet : la source sphérique émettait une impulsion d'une milliseconde, le micro récepteur était omnidirectionnel. A l'aide d'un micro directif, il est possible d'autre part de savoir de quelle direction provient telle réflexion, dans une certaine limite tout au moins.

Certains auteurs ont utilisé les échogrammes avec toutes les possibilités données par la méthode expérimentale, et en particulier la répartition spatiale de l'énergie incidente, illustrée par le "hérisson" de Meyer [3.29] [3.48] [3.49] [3.50] [3.51] . En soi, cette méthode est déjà riche d'enseignements, mais par la suite, on a essayé de tirer des échogrammes certains coefficients numériques particulièrement adaptés au problème de la perception auditive.
Certains d'entre eux permettent d'estimer la "qualité" d'une salle, c'est-à-dire la déformation apportée à une information acoustique, en grande partie d'ordre esthétique [3.52] [3.53] [3.54] [3.40] [4.6]
D'autres coefficients semblent particulièrement liés au problème de l'intelligibilité, et c'est à eux que nous nous attacherons en particulier.

L'exploitation des échogrammes repose en grande partie sur les travaux de Haas, qui est le premier chercheur à avoir étudié systématiquement le problème de l'intégration auditive des échos, sur lequel nous aurons l'occasion de revenir plus longuement.
Supposons qu'on écoute une phrase, prononcée en chambre sourde, et qu'à l'aide d'un dispositif particulier, on puisse juxtaposer au son direct une réflexion artificielle, décalée de millisecondes. Si les hauts parleurs qui émettent, l'un le son direct, l'autre le son réfléchi sont assez distants l'un de l'autre, tout en étant à la même distance de l'auditeur, on peut observer le phénomène suivant : tant que le délai τ est inférieur à un certain intervalle de l'ordre de 20 à 40 ms, les deux signaux se mélangent et sonnent comme un seul, sans effet de dédoublement, c'est-à-dire sans effet d'écho.
Bien entendu, cette propriété d'intégration auditive des échos est due à l'inertie de l'oreille. Nous reviendrons plus loin sur l'ensemble des mesures de perception des échos, qui constitue une partie très importante de notre travail.

D'autres travaux ont été effectués, parallèlement à ceux de Haas, vers les années 1950, pour essayer de tirer des renseignements substantiels des échogrammes. A la suite de Mason et Moir [3.56] Bolt et Doak ont proposé, en 1950, l'utilisation de courbes donnant le niveau d'un écho d'effet perturbateur constant en fonction de son retard [3.57] . Ces courbes devaient permettre de relever la présence d'échos nocifs sur un diagramme de réflexions.

b) La "netteté"de Thiele.-

Thiele compte parmi les premiers utilisateurs des échogrammes, mais il a en outre proposé l'emploi d'un coefficient appelé "netteté", qui est le rapport de l'énergie utile à l'énergie totale contenue dans l'échogramme [3.27]

Pour déterminer l'énergie utile, Thiele a choisi un délai maximal de 50 ms pour les réflexions qui suivent le son direct. Au-delà de ce délai, les réflexions sont supposées avoir un effet perturbateur, assimilables à un bruit qui masque le son utile. Si l'échogramme reçu est une pression qui varie en fonction du temps, soit $p(t)$, le son direct étant reçu à l'instant $t = 0$, la netteté est définie comme le rapport de l'énergie utile à l'énergie totale :

$$D = \frac{\int_0^{50\ ms} p^2(t)\,dt}{\int_0^{\infty} p^2(t)\,dt}$$

C'est donc un coefficient compris entre 0 et 1, qui peut caractériser l'acoustique d'une salle : Thiele a effectué de nombreuses mesures expérimentales, mentionnées dans l'article [3.27] ainsi que dans une publication ultérieure [3.29] . Thiele n'a pas utilisé ce coefficient pour chiffrer l'intelligibilité, mais seulement pour évaluer une caractéristique d'une salle. Bien entendu, certains auteurs et Januska en particulier ont utilisé cette netteté pour établir une prévision de l'intelligibilité [3.38] Nous reviendrons ultérieurement sur ce type de prévision.

c) Les travaux de Lochner et Burger.-

Un travail très important sur l'intelligibilité a été effectué par Lochner, Burger et Meffert à l'Université de Praetoria, en Afrique du Sud : les publications sont réparties entre les années 1958 et 1963 [3.30], [3.31] , [3.32] .

La méthode de Lochner est basée sur l'exploitation des échogrammes enregistrés expérimentalement. Pour évaluer l'énergie utile, Lochner a eu soin d'étudier les propriétés d'intégration auditive de l'oreille, avec la méthode suivante : dans une chambre sourde, deux haut-parleurs situés l'un à droite, l'autre à gauche de l'auditeur, avec des directions de 45° et -45° émettent respectivement le son direct et le son réfléchi : supposons que leur niveau soit identique, et tel que l'ensemble des deux haut-parleurs donne une intelligibilité de 50%, mesurée sur des listes de logatomes : soit N le niveau de chacun des haut-parleurs : si la voie retardée a un délai faible ou même nul, le son retardé est parfaitement intégré au son direct, et l'intelligibilité de 50% est celle donnée par un niveau total de N + 3 dB.
Si maintenant le son retardé l'est assez pour donner un effet d'écho, Lochner estime à juste titre que son intensité n'est pas complétement assimilée à celle du son direct. Pour obtenir une intelligibilité de 50%, il faudra que le son direct et l'écho aient une intensité N' ($>$ N). Lochner estime alors que l'écho augmente le niveau du son direct d'un nombre de décibels égal à N + 3 - N' $<$ 3. Par exemple, si N' = N + 1, l'écho n'aura renforcé le son direct que de 2 dB.

Nous ferons deux reproches à ces mesures d'intégration auditive : d'une part, on ne voit pas pourquoi le son direct et l'écho sont sensés provenir des directions $\pm$ 45°. Les mesures que nous avons effectuées, et qui font l'objet du chapitre V montrent que la situation est assez différente suivant les directions du son direct et de l'écho. D'autre part, la mesure directe de l'intégration auditive des échos à l'aide de mesures d'intelligibilité nous paraît assez artificielle et contestable.
En effet, Lochner en déduit un réseau de courbes qui donnent la fraction $\alpha(t)$ de l'énergie de l'écho intégrée au son direct. Cette fonction décroît de 1 à 0, et Lochner a trouvé que pour t = 95 ms, on avait pratiquement $\alpha(t) = 0$. (Fig. 3.15).

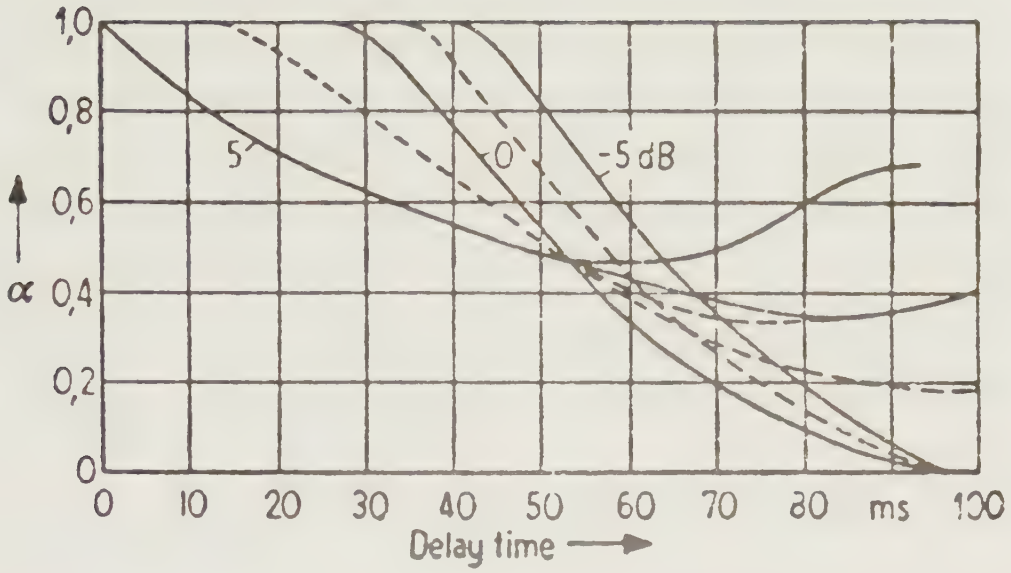


Fig. 3.15 - Courbe d'intégration auditive, suivant Lochner et Burger.-

L'énergie utile est alors :

$$E_u = \int_0^{95\ ms} \alpha\,(t)\ p^2(t)\ dt$$

L'énergie nuisible est évaluée suivant l'expression :

$$E_n = \int_{95ms}^{\infty} p^2(t)\ dt$$

Il y a quelque raison, à notre avis, de s'étonner qu'un écho arrivant avant 95 ms ne puisse pas avoir un effet perturbateur : on ne tient compte de son effet que par une contribution partielle à l'énergie utile.Nous aurons l'occasion, au Chapitre V, de revenir sur l'évaluation de Lochner.

Le reste de ce travail nous semble particulièrement convainquant : connaissant les valeurs de E_u et E_n , Lochner a déduit des mesures effectuées en chambre sourde la valeur intrinsèque de l'énergie utile. Il s'est alors reporté à un réseau de courbes qui donnent l'intelligibilité en fonction du niveau de la parole et d'un bruit de masque dont le spectre est identique à celui de la parole. On obtient alors une prévision d'intelligibilité qui semble donner des résultats en accord avec les mesures traditionnelles d'intelligibilité.

D'autre part, Lochner et ses collaborateurs n'ont pas envisagé l'effet directionnel de l'énergie incidente.

d) Travaux de Niese.-

Les travaux de Niese sur l'intelligibilité de la parole dans les salles ont été effectués entre 1955 et 1961, à l'Université de Dresde. Il s'agit d'un travail considérable, parallèle à celui de Lochner, et qui procède d'une façon générale analogue : on détermine expérimentalement les caractéristiques de l'oreille, et on en déduit sur l'échogramme une énergie utile et une énergie nuisible qui permettent de déterminer un "degré d'écho" (Echograd) [3.33] .

Ce degré d'écho est utilisé pour calculer un facteur de salle qu'on introduit dans la formule de Knudsen : cette détermination est effectuée en comparant les valeurs du degré d'écho et les mesures d'intelligibilité [3.34] . D'autre part, Niese s'est penché sur le problème particulier de la forme de l'impulsion émise: en effet, l'échogramme obtenu dépend du type de signal produit [3.35] . Niese a également envisagé quelle caractéristique de réception l'on devait donner au micro récepteur [3.36]

Les principaux résultats de ce travail sont rassemblés dans son article paru dans "Acustica" [3.37] . L'importance des travaux de Niese nous amène à rappeler quelques points essentiels. L'intégration auditive est étudiée de la façon suivante : on compare deux signaux acoustiques distants de 0,5 secondes : l'un est constitué par un train d'onde de 10 ms de

durée, et de niveau u_1 , l'autre est constitué par deux trains d'onde de 5 ms de durée pour chacun, et de niveau u_2 : la durée totale de ce deuxième signal est de T ms : on commence une mesure pour T = 10 ms, et l'on ajuste le niveau u_1 afin que les deux signaux aient la même sonie, ou sonorie, c'est-à-dire la même intensité sonore subjective. Bien entendu, on obtient $u_1 = u_2$. Lorsque T est supérieur à 15 ms, les valeurs de u_1 deviennent inférieures à u_2. Lorsque T est supérieur à 40 ms, les deux impulsions qui constituent le deuxième signal apparaissent séparées, il n'y a plus d'intégration. Des fréquences de 500 Hz, 1000 Hz et 3000 Hz ont été utilisées pour les trains d'onde sans donner de résultats différents. La courbe $u_1(T)$ permet de définir les deux limites d'intégration "partielle" des échos, c'est-à-dire 17 ms et 33 ms (Fig. 3.16)

Niese évalue d'autre part l'inertie de l'oreille en comparant la sonorie d'un train d'onde de durée T_2, de niveau p_2 avec celle d'un son stationnaire de niveau p_1 .
Pour diverses valeurs de T_2, on ajuste le niveau p_1 de telle sorte que la sonorie des deux signaux soit la même. On obtient alors une courbe dont l'expression approchée est :

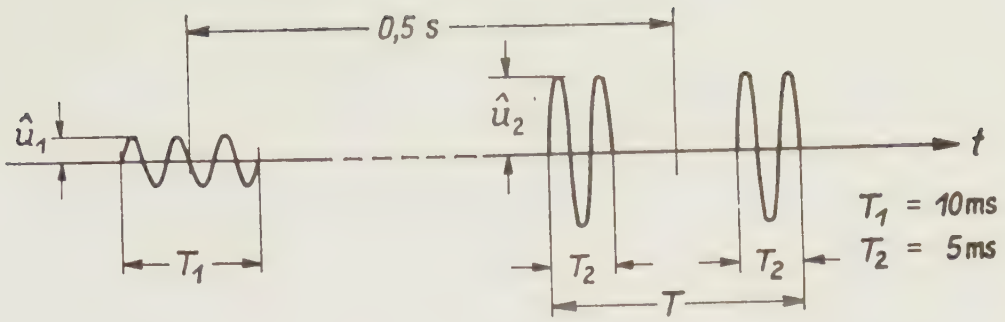
$$p_1 = p_2 \sqrt{1 - e^{-T_2/\tau_2}} \qquad \text{(Fig. 3.17)}$$

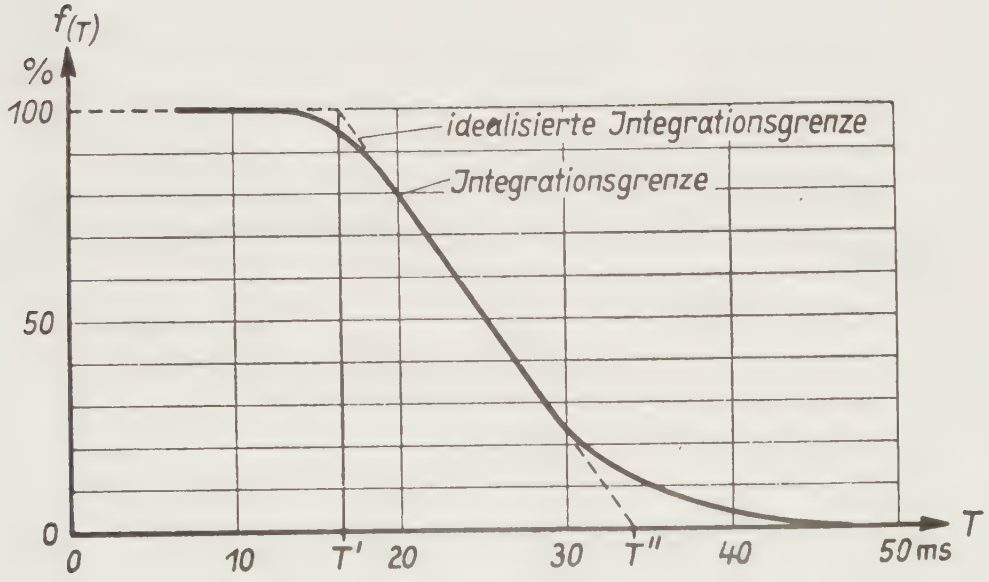
La constante τ_2 obtenue expérimentalement est la constante d'intégration de l'oreille, caractéristique de son inertie.
Niese utilise les constantes ainsi trouvées de la façon suivante :
l'énergie utile donnée par l'échogramme a pour expression :

$$E_u = \int_0^{t2} \left\{ |p(t)| \quad A(t) \right\}^2 \quad dt$$

où $A(t) = \begin{cases} 1, & \text{pour } 0 \leqslant t < t_1 \\ \dfrac{t - t_2}{t_1 - t_2} & " \quad t_1 \leqslant t \leqslant t_2 \\ 0 & " \quad t > t_2 \end{cases}$

avec t_1 = 17 ms et t_2 = 33 ms. (Fig. 3.18)





. . Courbe d'intégration des échos, selon Niese.

Le niveau u_2 est fixe, et l'on trace la courbe $f(T)=u_1/u_2$ en ajustant le niveau u_1 de sorte que la sonie des signaux de durée T_1 et T soit la même.

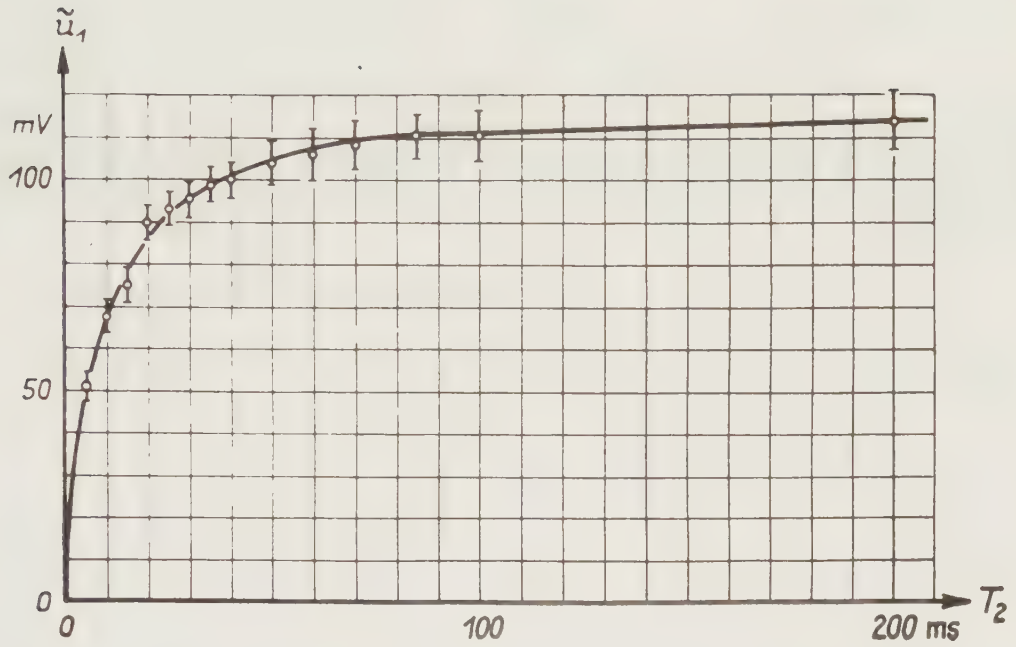


Fig. 3.17. Courbe d'intégration auditive de Niese.
On compare la sonie d'un son de niveau fixe et de durée T_2 à celle d'un son permanent de niveau u_1 : on ajuste ce niveau pour que la sonie des deux signaux soit la même.

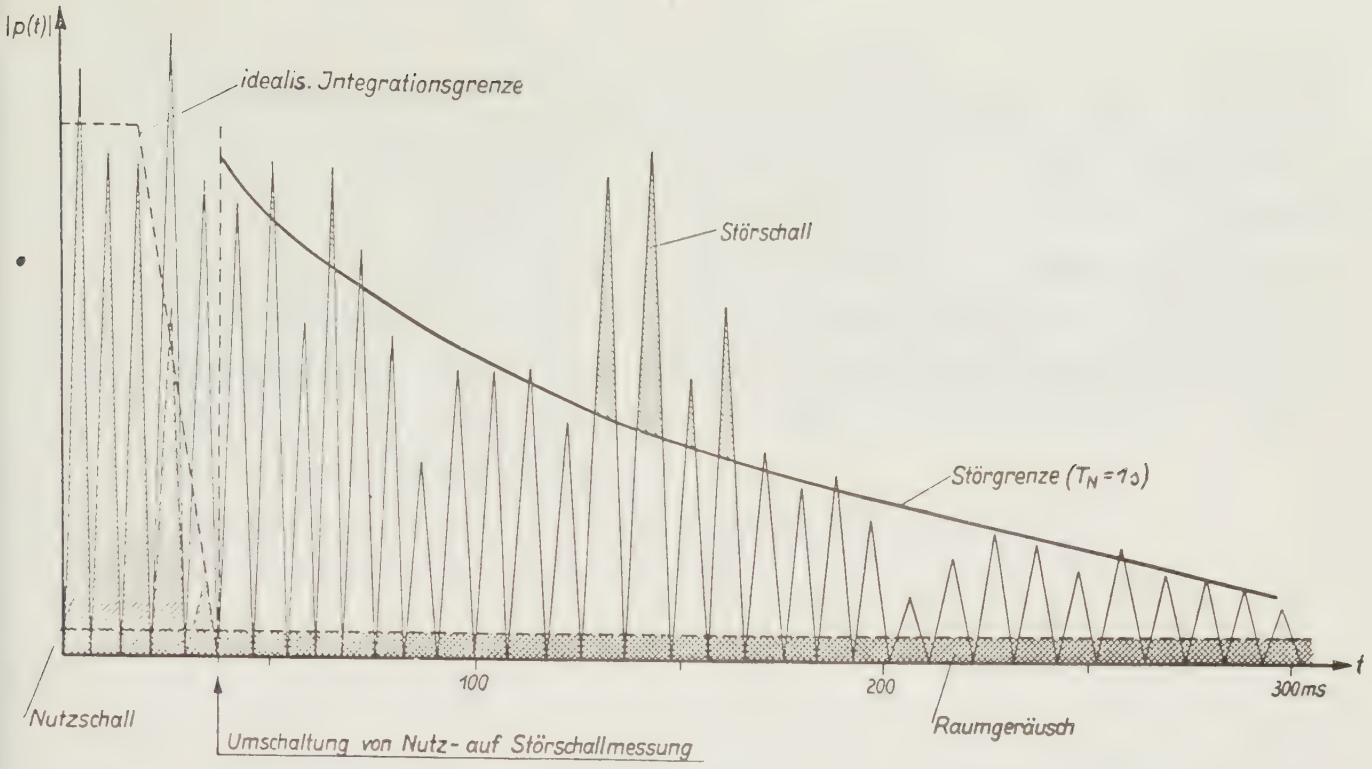


Fig. 3.18 - Définition de l'énergie nuisible d'après Niese [3.34.]

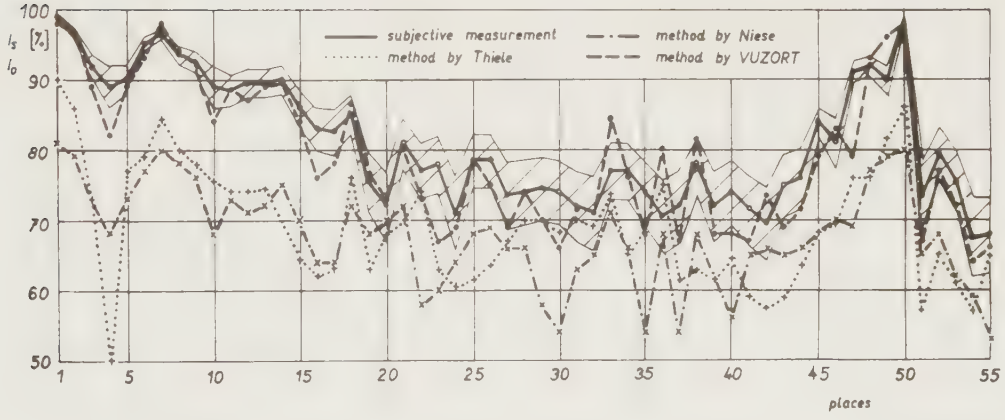


Fig. 3.19 - Mesures d'intelligibilité de Januska, et prévisions suivant différentes méthodes. En abcisse, on a porté le numéro des différentes places envisagées.-

Niese définit l'énergie nuisible en faisant abstraction de la réverbération: ainsi, une salle dont la réverbération serait parfaitement régulière aurait une énergie nuisible nulle : Niese fait intervenir la durée de réverbération dans la formule de prévision de l'intelligibilité en utilisant un coefficient particulier. L'énergie nuisible est donc l'énergie qui émerge d'une courbe de réverbération corrigée par la courbe de décroissance de la sonorie du système auditif : pour l'oreille, une salle qui aurait une constante de temps de 23 ms, et par conséquent une durée de réverbération de 317 ms doit paraître parfaitement sourde.
L'énergie nuisible est :

$$E_d = \int_{t_2}^{\infty} \left\{ |p(t)| - B(t) \right\}^2 dt \text{ , pour } |p(t)| \geqslant B(t)$$

$$\text{avec } B^2(t) = \left\{ 2 \int_0^{t_2} |p(t)|^2 dt \right\} \frac{\tau_R e^{-(t-t_2)/\tau_R} - \tau_0 e^{-(t-t_2)/\tau_0}}{\tau_R - \tau_0}$$

La valeur de B(t) pour $t = t_2$ est un niveau de crête moyen de la fonction p(t) entre 0 et t_2. τ_R est la constante de temps de la salle, dont la durée de réverbération est égale à $13{,}8\,\tau_R$.

Le degré d'écho ε est alors :

$$\varepsilon = E_d \; / \; (E_d + E_u)$$

Le coefficient ε sert alors à évaluer le facteur k_F qui intervient dans la formule proposée par Knudsen, et que Niese reprend.

$$It = 0{,}96\, k_E \,.k_B \,.k_R \,.k_F$$

Les travaux expérimentaux de Niese l'ont amené à prendre la relation suivante :

$$k_F = \sqrt[4]{1 - \varepsilon}$$

Il faut remarquer que Niese a pris soin d'utiliser une tête artificielle dont le diagramme de rayonnement est analogue à celui d'un locuteur, pour

émettre l'impulsion. D'autre part, le dispositif de réception est également constitué par une tête artificielle dotée de deux microphones en cardioïde qui simulent le diagramme de réception d'un opérateur humain. Nous verrons plus loin qu'à notre avis, l'utilisation d'un tel récepteur n'apporte guère d'amélioration dans la prévision de l'intelligibilité, pour la simple raison que le diagramme de réception des oreilles et l'influence de la direction d'un bruit de masque sont deux questions très différentes, sinon opposées.

En définitive, Niese a fourni sans doute le travail le plus complet à ce jour sur la prévision de l'intelligibilité dans les salles. Ce travail présente, à notre avis, les défauts suivants :
- Les courbes d'intégration de l'oreille sont basées sur des mesures dont la validité est contestable, dans le sens où un train d'onde ne schématise pas forcément la parole.
- D'autre part, on considère, comme dans les travaux précédents, chaque écho séparément, quant à son intégration au son direct, et indépendamment de ceux qui peuvent le précéder.
Niese n'a pas envisagé le problème de la direction des échos, et réduit le problème du récepteur à son diagramme de réception.

e) Travaux de Januska.-

Les travaux de Januska ont été effectués à Prague, à partir des années 1967. Ils y sont d'ailleurs actuellement poursuivis.
Januska reprend la formule de Knudsen, en introduisant comme facteur de forme de la salle la netteté de Thiele, ou bien un facteur déduit du degré d'écho de Niese. D'autre part, Januska considère que l'influence de la durée de réverbération est incluse dans le facteur k_F [3.38]
Si la netteté au sens de Thiele est $D = E_u/(E_u + E_n)$, avec

$$E_u = \int_0^{50ms} p^2(t)dt$$

$$E_n = \int_{50ms}^{\infty} p^2(t)dt$$

Januska a trouvé que la meilleure relation entre D et k_F était donnée par

la formule :

$$k_F = \sqrt[8]{D}$$

Le degré d'écho $\mathcal{E}'$ est calculé suivant la formule de Niese, mais en faisant intervenir dans l'évaluation de l'énergie nuisible, non pas la durée de réverbération réelle de la salle, mais une durée de réverbération optimale, variable suivant le volume de la salle envisagée. En fait, cette évaluation a pour effet d'inclure l'influence de la durée de réverbération, au cas où elle est supérieure à la durée optimale, dans l'évaluation de l'énergie nuisible. La relation choisie entre k_F et $\mathcal{E}'$, à l'aide des mesures expérimentales est :

$$k_F = \sqrt[8]{1 - \mathcal{E}'}$$

Januska réduit alors la formule de Knudsen à la suivante :

$$It = 100.k_E.k_F$$

estimant d'une part les auditeurs assez entraînés pour donner un résultat de 100% dans les meilleures conditions, et supposant nulle l'action du bruit extérieur.

Januska a effectué un travail expérimental considérable, en reprenant un matériel d'émission et de réception analogue à celui utilisé par Niese. Plus que d'une méthode originale, il s'agit d'un tour d'horizon et d'une confrontation de plusieurs techniques, alliée à un perfectionnement du procédé de Niese. La figure 3.19 donne un exemple des résultats de Januska obtenus pour une grande salle. Il faut également signaler le travail de H. KOCIC, en Pologne, qui constitue une série de mesures basées sur le procédé de Niese-Januska [3.39] .

f) Méthode de Kürer, et centre de gravité d'un échogramme.-

La distinction globale faite par plusieurs auteurs entre l'énergie utile et l'énergie nuisible, en considérant un échogramme, peut paraître sujette à caution, et nous aurons l'occasion de revenir longuement sur ce

problème. KÜrer a proposé, en 1969, l'utilisation d'un autre paramètre déduit d'un échogramme [4.41] . Il s'agit du "centre de gravité", ou moment du premier ordre, donné par l'expression suivante :

$$t_1 = \frac{\int_0^\infty t \ |p(t)| \ dt}{\int_0^\infty |p(t)| \ dt}$$

Pour tenir compte des différences qui peuvent apparaître entre les échogrammes enregistrés par les deux oreilles d'une tête artificielle, Kürer utilise de préférence à t_1 un centre de gravité binaural ainsi défini :

$$t_{1B} = \frac{\int_0^\infty t \ |p_D(t)| \ |p_G(t)| \ dt}{\int_0^\infty |p_D(t)| \ |p_G(t)| \ dt}$$

où $p_D(t)$ et $p_G(t)$ sont les échogrammes enregistrés à droite et à gauche, ayant les dimensions d'une pression sonore.

En reprenant la formule de Knudsen, Kürer remplace le produit $k_R \ k_F$ où k_R dépendait de la durée de réverbération et k_F de la forme de la salle, par un coefficient k_{1B} déduit de la valeur de t_{1B} , par la formule empirique :

$$k_{1B} = 1 - t_{1B}^2 \cdot 10^{-5} \qquad [4.42]$$

où t_{1B} est le centre de gravité binaural de l'échogramme, en millisecondes. Les résultats donnés par Kürer montrent une concordance assez satisfaisante entre l'intelligibilité et le coefficient k_{1B} : la dispersion est cependant assez forte et peut être de l'ordre de 10%. Il s'agit donc, à notre avis, d'une méthode globale, dont on ne saurait attendre des résultats plus précis que ceux donnés par Lochner ou Niese.
Il semble d'ailleurs que la méthode de Januska soit plus précise que celle de Kürer. D'autre part, on peut faire remarquer que le centre de gravité d'un échogramme n'a pas de signification physique évidente, et sa situation ne présente pas de lien apparent avec le processus d'intégration auditive des échos.

3.2.6. - Quelques autres travaux sur l'intelligibilité dans les salles.-

A côté des différents travaux que nous venons de mentionner, nous citerons quelques publications portant sur le même sujet, et dont certaines proposent une prévision de l'intelligibilité, ou apportent quelques remarques utiles en ce domaine.

Kurtovic [3.43] a présenté quelques mesures d'intelligibilité de la parole en présence de réverbération, fournie en l'occurence par une plaque : il apparaît nettement que la durée de réverbération est, en soi, un élément insuffisant pour permettre d'effectuer une prévision d'intelligibilité, et que c'est bien à la structure de l'échogramme qu'il faut faire appel pour avoir plus de renseignements.

Peutz et ses collaborateurs ont présenté lors de plusieurs communications [3.44] [3.45] [3.46] un travail expérimental important, d'où ils tirent les conclusions suivantes : en deçà d'une certaine distance D_c qui sépare le récepteur de la source, l'intelligibilité mesurée sur les consonnes décroît selon la loi suivante :

$$1 - It = (200\, D^2 \tau^2 / V + a)\ \%, \text{ où}$$

$$D < D_c = 0{,}20 \sqrt{V/\tau}$$

Au-delà de la distance D_c, l'intelligibilité reste constante :

$$1 - It = (9.\tau + a)\ \%$$

Les distances sont exprimées en mètres, le volume de la salle V en m^3, la durée de réverbération τ en secondes; a est un facteur de correction.

Ces formules empiriques ne peuvent pas donner de résultat très précis : le fait de considérer que l'intelligibilité reste constante au-delà d'une certaine distance est également approximatif, et peut être justifié par un calcul simple, comme celui que propose Franssen [4.47] si l'on considère l'énergie du son direct comme négligeable devant l'énergie réverbérée utile.

D'autre part, Peutz et Klein [4.46] ont effectué des mesures d'intelligibilité en présence d'un écho. Les renseignements donnés sont malheureusement très réduits : cependant, il apparaît qu'une réflexion, même proche du son direct, et retardée de 20 ms seulement, peut avoir un effet perturbateur. Une telle constatation ne pouvait que nous encourager à regarder de plus près le problème de la perception des échos.

3.3.- Conclusion.-

Nous avons essayé, au cours de ce chapitre, de faire le point sur les différentes techniques qui permettent de faire une estimation de l'intelligibilité de la parole dans une salle.

Nous ferons les remarques suivantes : les méthodes "sophistiquées", telles que la corrélation, la fonction de transfert de modulation nous paraissent trop éloignées de la réalité pour donner une prévision de l'intelligibilité vraiment précise. Ces méthodes négligent d'autre part délibérement le problème de l'audition binaurale, et on voit assez mal comment cet aspect des choses pourrait venir s'y greffer. D'autre part, ces méthodes restent assez éloignées des problèmes d'acoustique des salles. On se contente d'envisager un signal "détérioré" sans envisager les causes de sa détérioration. Le seul fait qu'on néglige souvent le rôle du niveau du signal reçu est significatif à cet égard.

Notre préférence est donc allée à des méthodes plus naturelles, qui reposent sur l'évaluation d'une énergie utile et d'une énergie nuisible. Les concepts qui permettent une approche du problème ne sont pas basés sur l'évaluation de l'énergie du son du son direct, mais sur celle de l'énergie associée au son direct, telle que l'ont envisagée Niese, ou Lochner. Cette évaluation repose sur l'examen des échogrammes. Jusqu'à ce jour, on avait utilisé uniquement des techniques expérimentales pour obtenir des échogrammes. Il nous a paru particulièrement intéressant de savoir s'il n'était pas possible d'obtenir ces échogrammes par un calcul, fût-il basé sur une théorie simpliste à certains égards. Ces différents calculs, appliqués au cas de plusieurs salles réelles ou fictives feront l'objet du chapitre IV.

A supposer que pour une salle donnée, on soit en mesure d'obtenir un tel échogramme, une question restait posée : comment déduire de cet échogramme une prévision de l'intelligibilité? A cette question, Niese ou Lochner ont apporté une réponse du plus haut intérêt : cependant, nous avons déjà eu l'occasion d'émettre des réserves sur la courbe d'intégration de l'oreille qu'ils ont utilisée pour évaluer l'énergie utile et l'énergie nuisible. D'autre part, aucun de ces auteurs n'a envisagé le rôle de la direction des

échos. Ceci nous a conduit à la réalisation d'un "Champ Sonore Synthétique" qui permet de démonter le mécanisme de la perception des échos.

Il apparaît que, depuis les travaux déjà anciens de Haas, aucun auteur n'a beaucoup utilisé un tel montage pour étudier ce phénomène. Nous ne connaissons à ce sujet que des travaux limités que nous aurons l'occasion de citer plus loin.

L'exploitation du champ sonore synthétique fera l'objet du Chapitre V, au terme duquel nous pourrons faire le point sur la contribution originale que nous avons pu apporter au problème de l'intelligibilité de la parole dans les salles.

x x x

B I B L I O G R A P H I E

3.1. SADEK (D.).- De l'intelligibilité (J.F.O.R.L., XI, 6, pp. 889-905, 1962).-

3.2. CHAVASSE (P.).- De la notion d'intelligibilité (J.F.O.R.L., XI, 6, pp. 857-873, 1962).

3.3. CHAVASSE (P.).- Phonétique statistique de la langue française.- (Ann. des télécomm., n° 1, pp. 5-16, 1948).-

3.4 JANUSKA (I.), SPRONCK (F.), DENDAL (J.).- Mesure de l'intelligibilité syllabique par une méthode de libre choix.- (Electroacoustique Belge, n° 21, pp.2-15, 1972).-

3.5 HECKER (M.L.), STEVENS (K.N.), WILLIAMS (C.E.).- Measurements of reaction time in intelligibility tests.- (JASA, 39, 6, pp. 1188-9, 1966).-

3.6 POLLACK (I.).- Message probability and message reception.- (JASA, 36, 5, pp. 937-45, 1964).-

3.7 STUCKEY (C.W.).- Investigation of the precision of an articulation testing programm.- (JASA, 35, 11, pp. 1782-7, 1963).-

3.8 CANAC (F.).- L'acoustique des théâtres antiques, ses enseignements.- (Ed. du CNRS, 1968).-

3.9 CHAVASSE (P.).- L'application des moyens d'analyse de la qualité des transmissions téléphoniques.- (Ann. des télécomm., cahiers d'acoustique, 24, pp. 109-22, 1950).-

3.10 LEHMANN (R.).- Etude psychophysique de l'intelligibilité du langage.- (Ann. des télécomm. 17, n° 9-10, Sept-Oct 1962, pp. 242-303).-

3.11 JANUSKA (I.).- Application of the impulse measuring method for objective speech intelligibility prediction in enclosures.- (Electroacoustique belge, n° 13, pp. 5-22, 1968).-

3.12 KRYTER (K.).- Projet de normalisation pour les tests vocaux.- (AFNOR S 30/GN, doc. 9), Sept 1973).-

3.13 MATRAS (J.J.).- Acoustique et électroacoustique, Ed. Eyrolles, Paris.-

3.14 KNUDSEN (V.).- Architectural acoustics.- (J. Wiley and sons, New-York, 1950).-

3.15 FRENCH (N.R.), STEINBERG (J.C.).- Factors governing the intelligibility of speech sounds.- (JASA, 19, 1, pp. 90-119, 1947).-

3.16 BOLT (R.H.), MAC DONALD (A.D.).- Theory of speech masking by reverberation (JASA, V 21, 6, pp. 577-80, 1949).-

3.17 KORN (T.).- Théorie générale de l'intelligibilité dans les salles.- (Cahiers d'acoustique, n° 20, pp. 89-93, 1950).-

3.18 JANSSEN (J.H.).- A method for the calculation of speech intelligibility under conditions of reverberation and noise.- (Acustica, V 7, pp. 305-10, 1957).-

3.19 MAX (J.).- Traitement du signal, Masson ed., pp. 35-36, 1972.-

3.20 FOURDOUEV (V.).- Evaluation objective de l'acoustique des salles.- 5° Congrès International d'Acoustique, Liège, 1965.-

3.21 BLADIER (B.) et SANTON (F.).- Sur la déformation de la parole dans des conditions de réverbération.- Note aux C.R. Acad. Sc., t. 266, pp. 367-9, 12 fev. 1968.-

3.22 SANTON (F.).- Application d'une méthode de corrélation à l'intelligibilité de la parole dans les salles.- D.E.S., Fac. des Sc. Aix-Marseille, Oct. 1968.-

3.23 LEIPP (E.).- Traitement optique de l'information acoustique.- Revue d'acoustique, n° 26, pp. 182-9, 1973.-

3.24 RISSET (J.C.).- Sur l'analyse, la synthèse des sons.- Thèse, Fac. Sc. Orsay, Mai 1967.-

3.25 HOUTGAST (T.) et STEENEKEN (H.J.M.).- The modulation transfer function in room acoustics as a predictor of speech intelligibility.- Acustica, V 28, pp. 66-73, 1973.-

3.26 BLADIER (B.).- Sur les réponses transitoires en acoustique des salles.- Acustica, 18, 6, pp. 309-22, 1967.-

3.27 THIELE (R.).- Richtungverteilung und Zeitfolgender Schallrückwürfe in Räumen.- Acustica, V 3, pp. 291-302, 1953.-

3.28 HAAS (H.).- Über den Einfluss eines Einfachechos auf die Hörsamkeit von Sprache.- Acustica VI, n° 2, pp. 49-51, 1951.-

3.29 MEYER (E.) et THIELE (R.).- Raumakustische Untersuchungen in zahreichen Konzertsälen und Rundfunkstudios unter Anwendung neuerer Messverfahren.- Acustische Beihefte (Acustica), V 6, n° 2, pp. 425-44, 1956.-

3.30 LOCHNER (J.A.P.) et BURGER (J.F.).- The subjective masking of short time delayed echoes by their primary sounds and their contribution to the intelligibility of speech.- Acustica V 8, n° 1, pp. 1-10 , 1958.

3.31 LOCHNER (J.A.P.) et BURGER (J.F.).- The intelligibility of speech under reverberant conditions.- Acustica, V 11, pp. 195-200, 1961.-

3.32 LOCHNER (J.A.P.) et MEFFERT (P.).- Analysis of reflexion patterns in Auditoria.- JASA, V 35, n° 9, pp. 1429-1431, 1963.-

3.33 NIESE (H.).- Vorschlag für die Definition und Messung der Deutlichkeit nach subjektiven Grundlagen.- (Hochfrequenztechnik und Electroakustik, V 65, n° 1, pp. 4-15, 1956).-

3.34 NIESE (H.).- Due prüfung der aurmakustischen "Echograd-kriteriums" mit Hilfe von Silbenverständlichkeitsmessungen.- Hochfreq und Electroak, V 66, n° 3, pp. 70-83, 1957.-

3.35 NIESE (H.).- Untersuschung über die Knallform bei raumakustischen Impulsmessungen.- Hochfrequenztechnik und Electroakustik, V 65, n° 3, pp. 98-108, 1956.-

3.36 NIESE (H.).- Untersuchung für die Richtcharäkteristik des Aufnahmen mikrophons bei raumakustischen Impulsmessungen.- Hochfreq. und Electroa, V 65, n° 6, pp. 192-200, 1956.-

3.37 NIESE (H.).- Die M essung der Nützschall une Echogradverteilung zur Beurteilung der Hörsamkeit in Raumen.- Acustica, V 11, pp. 207-213, 1961.-

3.38 JANUSKA (I.).- Application of the impulse measuring method for objective speech intelligibility in enclosures.- Electroacoustique belge, n° 13, pp. 5-22, 1968.-

3.39 JOCIC (H.).- Investigations on the speech intelligibility in large rooms with regard to the Polish language.- Symposium "Intelligibilité de la parole" - Liège 12-15 Nov. 1973.-

3.40 KÜRER (R.).- Zur Gewwinnung von Einzahlkriterien bei Impulsmessungen in der Raumakustik.- Acustica, V 21, n° 8, pp. 370-72, 1969.-

3.41 KÜRER (R.).- Einfaches Messverfahren zur Bestimmung der "Schwerpunktzeit" raumakustischen Impulsantworten.- 7° CIA, Budapest, Communication 23 A5, 1971.-

3.42 KÜRER (R.).- Studies on parameters determining the intelligibility of speech in auditoria.- Symposium "Intelligibilité de la parole", pp. 108-117, Liège, Nov. 1973.-

3.43 KURTOVIC (H.S.).- Contribution à la recherche de l'influence du son réfléchi sur l'intelligibilité.- 5° CIA, communication G 21.-

3.44 PEUTZ (V.M.A.) et TOLK (J.).- Speech intelligibility in reverberant rooms.- 5° CIA, communication A 52.-

3.45 PEUTZ (V.M.A.).- Articulation loss of consonants as a criterion for speech transmission in rooms.- Journal of the Audio Engineering Society of America, V 19, n° 11, pp. 915-919, Dec. 1971.-

3.46 PEUTZ (V.M.A.) et KLEIN (W.).- Articulation loss of consonants influenced by noise, reverberation and echo.- Symposium "Intelligibilité de la parole", Liège, Nov. 1973.-

3.47 FRANSSEN (N.V.).- L'intelligibilité naturelle dans les salles et celle améliorée par des moyens électroacoustiques.-

3.48 MEYER (I.).- Definition and diffusion in rooms.- JASA, V 26, N 5, pp. 630-36, 1954.-

3.49 JUNIUS (W).- Raumakustische Untersuchungen mit neueren Messverfahren in der Liederhalle Stuttgart.- Acustica, V 9, N 4, pp. 289-303, 1959.-

3.50 SCHODDER (G.R.).- Über die Verteilungen energiereicherer Schallrückwürfe in Sälen.- Acustica, V 6, pp. 445-460, 1956.-

3.51 MEYER (I.).et BURGTORF (W.).- Über die Zeitabhängigkeit der Schaarichtungsverteilung in Räumen bei impulsartiger Anregung.- Acustica, V 7, pp. 313-19, 1957.-

3.52 REICHARD (W.).- Der Impuls-Schalltest une seine raumakustische Beurteilung.- 6° ICA, Communication GP-2-2.-

3.53 REICHARD (W.).- Bedeutung und Auswertung der Anfangsreflexion in Reflektogrammen für die Raumakustische Beurteilung von Musik.- 7° ICA, Communication 24 S 18.-

3.54 REICHARD (W.) et KUSSEV (A.).- Prüfung der Balance in Konzerträumen bei impulsiver Erregung.- Acustica, V 31, N 1, pp. 21-34, 1974.-

3.55 MASON (C.A.) et MOIR (J.).- Acoustics of cinema auditoria.- J. Inst. Elec. eng., 88, III, N 3, p. 175, 1941.-

3.56 BOLT (R.H.) et DOAK (P.E.).- A tentative cirterion for short term transient response in auditoriums.- JASA, V 22, N 4, pp. 507-509, 1950.-

X
X X
X

4.- CALCUL DES ECHOGRAMMES DANS UNE SALLE -

Nous avons déjà eu l'occasion de parler longuement des échogrammes et de leur utilisation. L'enregistrement des échogrammes a été souvent pratiqué. Déduire une prévision d'intelligibilité de mesures physiques effectuées dans une salle existante constitue un premier pas, et permet d'éviter des auditions de logatomes, toujours fastidieuses. Mais la meilleure solution, dans le cas d'une salle à édifier est de prévoir à l'avance si l'intelligibilité y sera suffisante : autrement dit, il vaut mieux ne pas avoir à construire une salle pour se rendre compte que sa qualité acoustique laisse à désirer, même si des corrections seraient possibles a posteriori.

On sait que lorsqu'un architecte fait un projet de salle de spectacle, l'acousticien consulté commence souvent par étudier les propriétés de cette salle à l'aide d'un modèle réduit, ou maquette, dont les caractéristiques physiques seront connues grâce à l'utilisation de longueurs d'ondes elles aussi réduites. Le passage des fréquences acoustiques aux ultrasons n'est pas sans poser certains problèmes délicats, en particulier en ce qui concerne le rôle du coefficient d'absorption des matériaux qui tapissent la salle, et la maquette. Plutôt que de construire un modèle réduit, il peut paraître séduisant d'obtenir des renseignements substantiels sur la salle à édifier, par une technique numérique. Nous allons étudier quelques moyens théoriques et pratiques d'obtenir la "réponse" d'une salle à un signal de forme quelconque.

4.1.- Mise en équation du problème.-

Il s'agit d'un problème de diffraction d'un signal quelconque par les parois d'une enceinte où il se trouve émis. La propagation de ce signal est régie par l'équation de d'Alembert. Si $\varphi(M,t)$ est le potentiel des vitesses attaché à un point M situé dans une enceinte V, à un ins-

tant t, $\varphi(M,t)$ est lié au terme source $f(M_0, t)$ par la relation :

$$\left(\Delta - \frac{\partial^2}{c^2}\right)\varphi(M,t) = f(M_0,t)$$

$f(M,t)$ est nul si M n'est pas en M_0, c'est-à-dire si M n'a pas la situation de la source. Sur la frontière de l'enceinte considérée, c'est-à-dire sur les parois de la salle, le potentiel des vitesses est soumis aux conditions aux limites de la forme :

$$(\partial_n + \alpha \partial_t)\varphi(P,t) = 0$$

si P est un point de la frontière, et n la normale à cette surface, extérieure à l'enceinte. α est l'admittance de la paroi.

En partant de ces deux équations, le problème est difficile à traiter directement, car il faut faire appel à des fonctions de Green que l'on ne peut pas expliciter.
Une méthode envisageable est celle de la transformation de Fourier.

4.2.- Calcul au moyen de la transformation de Fourier.-

Envisageons les transformées de Fourier des fonctions $\varphi(M,t)$ et $f(M_0,t)$ soit :

$$\Phi(M,\omega) = \mathcal{F}\ \varphi(M,t)$$

$$F(M_0,\omega) = \mathcal{F}\ f(M_0,t)$$

L'équation de propagation, relative à une fréquence ω est alors l'équation de Helmholtz :

$$(\Delta + \omega^2/c^2)\,\Phi(M,\omega) = F(M_0,\omega)$$

La condition aux limites devient alors:

$$(\partial_n + j\omega\,\alpha(\omega))\ \Phi(P,\omega) = 0$$

Posons $\Phi = \Phi_0 + \Phi_1$ où Φ_0 est le champ incident :

$$\Phi_0 = \int F(M_0, \omega)\, g(M, M_0)\, dM_0$$

$g(M,M_0)$ est la singularité fondamentale associée à l'opérateur $\Delta + \omega^2/c^2$

$$g(M,M) = - e^{ikr(M,M_0)}/r(M,M_0)$$

Φ_1 est le champ diffracté que l'on peut prendre sous la forme d'un potentiel de simple couche :

$$\Phi_1(M) = \int_\Sigma \mu(P)\, g(M,P)\, dS(P)$$

où dS est un élément de la surface Σ .

En tenant compte de la discontinuité du gradient de Φ_1 au passage de la frontière Σ , et en reportant les valeurs de Φ_0 et Φ_1 dans la condition aux limites, on obtient l'équation intégrale suivante :

$$-\frac{\mu(P)}{2} + \int_\Sigma \mu(P')\,[\partial_n(P) + j\omega\alpha(P)]\, g(P,P')\, dS(P') = -[\partial_n(P) + j\omega\alpha(P)]\, \Phi_0(P)$$

La résolution de cette équation intégrale en $\mu(P)$ permet de connaître l'expression du champ Φ_M , pour chaque valeur de ω.

Si ce travail est effectué pour un grand nombre de fréquences ω , on pourra alors passer à la transformée de Fourier inverse :

$$\varphi(M,t) = \mathcal{F}^{-1}\Phi(M,\omega)$$

et obtenir la fonction désirée. Rappelons que les "observables", c'est-à-dire les valeurs de la pression sonore et de la vitesse vibratoire se déduisent respectivement de $\varphi(M,t)$ par les expressions suivantes :

$$p = \rho_e\ R_e\ (\partial\Phi/\partial t)$$

$$\vec{v} = R_e\ (-\overrightarrow{\text{grad}}\,\Phi)$$

où ρ_e est la masse spécifique de l'air, et R_e désigne la valeur réelle des fonctions envisagées.

Dans le cas d'une géométrie particulièrement simple, on peut déterminer facilement $\Phi(M,\omega)$, sans avoir même à faire appel à une technique de potentiels de couche, comme l'ont montré en particulier les travaux de Morse et Bolt, [4.2] , Morse et Ingard [4.3] , ou ceux plus récents de Doak et Dodd [4.4] . Le passage à une transformation de Fourier ne présente pas de difficultés particulières, il est parfaitement envisageable dans la mesure où le calcul de la fonction $\Phi(M,\omega)$ échantillonnée sur un nombre de fréquences suffisant est rapide.

Mais le cas de la salle rectangulaire, sur lequel nous aurons l'occasion de revenir, ne présente qu'un intérêt limité. Il présente l'avantage d'être facilement abordable, et de fournir sans effort l'ensemble des fréquences propres par une seule formule, celle des cordes vibrantes généralisées à trois dimensions. Certains font remarquer que le cas d'une salle rectangulaire est moins scolaire ou artificiel qu'on veut bien le dire [4.5] , nombre de salles ayant le bon goût de présenter cette géométrie, de la salle de séjour d'un appartement jusqu'aux fameuses salles de concert de Vienne ou de Boston, si toutefois, on veut bien faire abstraction de la présence des balcons [4.6].

Cependant, il faut bien admettre que seule une méthode théoriquement applicable à une configuration de salle tout-à-fait quelconque peut satisfaire l'esprit. A notre connaissance, seule la méthode utilisée par G.P. Extrémet [4.7] [4.8] pour l'ellipsoïde de révolution, puis pour une salle heptaédrique [4.9] donne une solution générale au problème de la "réponse" d'une salle à une excitation en régime permanent.
A partir de ce résultat, peut-on envisager un passage au cas d'une excitation transitoire, au moyen de la transformation de Fourier ?
Pour des raisons pratiques, cela semble difficile : d'une part, le calcul de $\Phi(M,\omega)$ est déjà passablement laborieux, pour une seule fréquence .
Le volume des calculs risquerait fort d'être prohibitif si l'on voulait obtenir une courbe $\Phi(M,\omega)$ pour suffisamment de valeurs distinctes de ω.
D'autre part, l'exemple d'une salle réelle, celui de la salle heptaédrique que nous utiliserons également plus tard, a montré que les capacités de l'ordinateur obligeaient à limiter la valeur maximale des fréquences envisagées : autrement dit, il n'est pas possible, à l'heure actuelle, et avec

les possibilités de calcul que nous avons à notre disposition, d'envisager l'obtention de la fonction $\Phi(M,\omega)$ pour des valeurs de la fréquence $f = \omega/2\pi$, supérieures à 500 Hz. Indépendamment de toute considération de rentabilité des calculs, cet élément rend l'utilisation de cette méthode pour le calcul de la réponse transitoire d'une salle encore très problématique.

D'autre part, nous avons déjà eu l'occasion de mentionner que nous désirions obtenir un échogramme spatio-temporel, et non un échogramme temporel seulement : c'est dire que nous ne nous intéressons pas seulement à la succession dans le temps des différents échos qui suivent le son direct, mais également à leurs directions respectives. Or, la méthode que nous avons envisagée ci-dessus ne nous donnerait en tout état de cause qu'un échogramme temporel, ce que nous jugeons insuffisant pour le but que nous nous sommes proposés.

Existe-t-il alors une méthode "exacte" capable de donner tous les renseignements que nous désirons obtenir d'un échogramme spatio-temporel? A l'heure actuelle, ce domaine semble encore en friche. Nous exprimons sincérement le voeu qu'un procédé rigoureux puisse être développé.
Comme l'obtention par le calcul d'un échogramme spatio-temporel, aussi exact fut-il, ne marquait pas le bout de nos peines, puisque nous nous proposions de l'exploiter pour obtenir une prévision de l'intelligibilité, nous ne nous sommes pas engagés dans cette voie par trop incertaine.

La méthode que nous avons utilisée est à la base, sommaire, schématique, voire inexacte : mais elle permet sûrement d'atteindre l'objectif cherché. Nous allons exposer quel est le principe de cette méthode et comment on peut envisager le calcul numérique d'un échogramme.

4.3.- Procédés de calcul d'un échogramme au moyen de l'acoustique géométrique.-

La technique utilisée consiste à délaisser la rigueur liée à la considération du caractère ondulatoire du signal sonore, pour obtenir plus directement des renseignements suffisamment approchés concernant la réponse aux signaux transitoires.

Cette façon d'envisager les choses ne date pas d'hier : elle est très ancienne, et l'on croirait volontiers, avec F. Canac [3.8] , que les architectes de l'antiquité connaissaient les principes de l'acoustique géométrique. Bien entendu, cette méthode néglige délibérément les effets de la diffraction, donc il serait illusoire de vouloir obtenir un échogramme parfaitement conforme à la réalité expérimentale. Mais telle n'était pas notre prétention : nous voulions obtenir un échogramme assez proche de la réalité pour les renseignements que nous voulions en tirer, à savoir une prévision de l'intelligibilité de la parole dans les salles. Après avoir exposé les résultats relatifs à la vérification expérimentale, nous reviendrons sur les problèmes de l'exactitude de l'acoustique géométrique.

Nous aurons l'occasion de développer deux techniques : l'une utilise les images de la source par rapport aux différentes parois de la salle; l'autre une discrétisation du rayonnement sonore de la source sous forme de "rayons". La méthode des images a été utilisée en particulier par F. Canac [4.10] [3.8] dans ses études sur les théâtres antiques, elle était également mentionnée par Cremer [4.11] . Bolt, Doak et Westervelt [4.12] ont également utilisé les images, dans le cas d'une salle rectangulaire, pour étudier la réponse d'une salle à une impulsion. C'est dans le même but que Holtsmark les invoque également, toujours dans le cas d'une salle rectangulaire [4.13] [4.14] .

Tous ces travaux utilisaient, soit des moyens de calcul artisanaux, soit des formules globales pour prendre en considération l'effet des sources images. En fait, un nouvel essor de l'acoustique géométrique a été provoqué par l'utilisation des calculateurs numériques à hautes performances. Parmi les travaux récents qui ont été effectués en utilisant la méthode des images, nous citerons ceux de Pujolle [4.15] ,[4.16] , [4.17] et Tho Tuong [4.18] , relatifs aux problèmes de la réverbération, du libre parcours moyen, etc, dans le cas d'une salle rectangulaire, ainsi que ceux de Barron [4.19] , également relatifs à une salle de même géométrie.

Par rapport à ces travaux, le notre offre l'originalité du cas d'une salle polyédrique irrégulière traitée par la méthode des images. L'extension du problème ainsi envisagée nous semble digné d'intérêt. Quant à la deuxième méthode d'acoustique géométrique que nous avons utilisée, celle des rayons, son développement ne date que de l'utilisation des calculateurs numériques puissants. En effet, il n'est guère possible de suivre

l'histoire des rayons sonores sans avoir recours aux ordinateurs, alors qu'on peut invoquer, tout au moins dans le cas d'une salle rectangulaire, une formule qui donne la situation des sources images, et leur rayonnement. Schroeder semble avoir été l'un des premiers à utiliser cette méthode des rayons, dans le cas d'un modèle à deux dimensions (modèle plan) [4.20] Parallélement à Schroeder, les norvégiens Krokdtad, Ström et Sördal [4.21] ont utilisé cette technique dans le cas d'une salle réélle à trois dimensions (salle rectangulaire et salle en éventail). Plus récemment, Kuttruff [4.22] a utilisé une méthode de rayons pour étudier l'effet de diffuseurs placés dans une salle rectangulaire, ainsi que l'effet d'une réflexion diffuse sur la décroissance du son dans une salle. Nous citerons encore le projet d'étude de P. Armand [4.23] qui se situait dans la même perspective.

La contribution que nous apportons dans ce domaine n'est pas nouvelle, quant à son principe. Remarquons que, si c'est l'échogramme qui nous intéresse, les auteurs susnommés se sont attachés au problème de la décroissance sonore, mais ce problème-ci et celui-là ne relèvent pas de techniques de calcul différentes.

4.3.1.- Objet de l'étude.-

Le problème préliminaire est purement énergétique: étant donné une salle contenant une source d'énergie sonore et un récepteur, quelle est la fraction d'énergie recueillie au récepteur? Le problème principal résulte de l'adjonction de l'aspect temporel à l'aspect énergétique précédent : quelle est la répartition temporelle de l'énergie recueillie?

Le problème est formulé pour une salle, mais il s'étend aisément à une construction ouverte à l'air libre; il suffit que soit admise l'équivalence entre une paroi parfaitement absorbante et une ouverture, pourvu, évidemment, qu'aucune énergie sonore ne puisse arriver de l'extérieur par cette ouverture.

Nous n'apporterons aucune restriction de principe quant à la forme de la salle. Les rayons sonores étant supposés subir, au contact des parois, une réflexion suivant les lois de l'optique géométrique, il faudra que soient définis, en tout point de la paroi, la normale à celle-ci et le facteur de réflexion, fonction, éventuellement, de l'angle d'incidence.

Les restrictions porteront sur la source et le récepteur. La source, isotrope ou non, sera supposée ponctuelle. Le récepteur pourra être soit une portion de plan faisant partie des parois de la salle, soit une sphère située à l'intérieur de la salle. Nous le supposerons parfaitement absorbant, mais l'hypothèse contraire est également admissible.

Le milieu propagateur, l'air, sera supposé homogène et immobile. L'aspect temporel, pour important qu'il soit dans le but de notre étude, n'est pas essentiel quant à la méthode : il est fondé sur un élément, la vitesse de propagation, qui, dans les études d'optique géométrique n'est jamais à prendre en considération. Nous ne l'introduirons qu'après avoir exposé les deux méthodes provisoirement limitées au traitement de l'aspect énergétique.

4.3.2.- Aspect énergétique.-

Tant qu'il s'agira de l'aspect purement énergétique, nous trouverons quelque intérêt à pousser assez loin l'analogie avec le phénomène lumineux. Nous pouvons imaginer une salle dont les parois, ainsi que les détails éventuels seraient réfléchissants en tout point, le facteur de réflexion pouvant varier selon la position du point et l'angle d'incidence. Une source lumineuse ponctuelle, isotrope ou non, serait placée dans cette salle et l'on étudierait l'éclairement d'une portion délimitée de la paroi, ou encore d'une sphère située dans la salle. Cette analogie nous donne une image intuitive, quoique inexacte à certains points de vue, de la répartition du champ sonore dans la salle et sur le récepteur.

Pour définir les directions des rayons émis par la source S, nous attacherons au point S un système de coordonnées sphériques θ , longitude et φ , latitude. L'élément différentiel d'angle solide est :

$$d\omega = \cos\varphi \; d\theta \; d\varphi$$

$d\omega$ sera aussi considéré comme l'élément d'aire de la sphère trogonométrique centrée en S.

L'intensité d'émission I dans une direction $D(\theta,\varphi)$ est définie par le quotient différentiel $dW/d\omega$ de l'énergie émise par l'angle solide correspondant. Elle est généralement variable selon la direction.

Nous supposerons connue la fonction $I(\theta,\varphi)$. On a :

$$dW = I(\theta,\varphi) \cos\varphi \, d\theta \, d\varphi$$

Le rayon transportant au départ l'élément d'énergie dW, une fois émis dans la direction D, va suivre sa "destinée" avant de tomber sur le récepteur avec l'énergie :

$$dw = a \, d \, W$$

Le facteur a, "facteur d'affaiblissement", est égal à l'unité pour le cas d'un rayon allant directement de la source au récepteur. Plus généralement, il est égal à un produit de facteurs de réflexion. Il peut être nul, pour certains rayons, si les hypothèses concernant les parois de la salle permettent que ces rayons soient complétement absorbés avant d'arriver au récepteur.

S'il faut définir plus précisément l'expression "destinée" utilisée ici, disons : c'est une classe d'équivalence, deux rayons étant équivalents s'ils subissent le même nombre de réflexions entre la source et le récepteur et si les réflexions de même rang se font en deux points d'un même domaine des parois de la salle, ce domaine étant connexe et sans arêtes. Si l'on adjoint à cette dernière condition qui impose la continuité de la normale à la paroi, la condition de continuité du facteur de réflexion, on définit un autre ensemble de classes qui pourrait ailleurs, servir à définir différemment la "destinée."

La "destinée" d'un rayon qui tombe pour se réfléchir sur une arête de la salle n'est pas définie. On peut aussi bien dire qu'il a deux destinées distinctes, d'où, en général, deux valeurs différentes de a. Dans l'ensemble des directions, la facteur a qui est fonction de D, est généralement affecté de nombreuses discontinuités.

Imaginons ce qui se passe du point de vue des éclairements au niveau du récepteur, celui-ci étant considéré comme ayant des dimensions finies, ce qui est conforme à la réalité physique. Chacune des discontinuités du facteur a se traduit par une ligne de discontinuité de l'éclairement du récepteur. Celui-ci doit alors être imaginé comme strié d'un grand nombre de telles lignes, ce qui lui donne l'aspect d'une mosaïque (toujours, bien entendu, dans l'analogie que nous avons choisie pour avoir

une représentation visuelle du phénomène acoustique à l'étude).

Le problème à résoudre présente alors deux options :

a) calculer l'éclairement énergétique moyen du récepteur, c'est-à-dire le quotient w/A de l'énergie reçue par l'aire du récepteur;

b) calculer la limite du quotient ci-dessus lorsque les dimensions du récepteur sont réduites de façon telle qu'aucune discontinuité de l'éclairement ne subsiste à sa surface, le récepteur demeurant centré au point C où l'on désire ce résultat.
Pour tenir compte, en toute rigueur, de ce que le rapetissement du récepteur, qui est absorbant, peut modifier la répartition de l'énergie dans la salle, nous dissocierons la fonction "réception" de la fonction "absorption" du récepteur en définissant une variante de l'option précédente :

b') calculer la limite définie en b, mais en supposant que, pour l'absorption, le récepteur conserve ses dimensions. Cette limite possède une signification physique dans le cas du récepteur plan : c'est l'éclairement au point C.

Dans le cas général d'une salle présentant une ou plusieurs parois courbes, la notion de "destinée" ne perd pas sa signification, mais l'existence de telles parois occasionne des concentrations irrégulières de rayons appartenant à une même classe. Cet effet, se superposant à l'effet de mosaïque dontnous avons parlé, ôte à peu près toute signification à la notion d'éclairement local. Pour une telle salle, le problème se réduira alors à l'option a) .

Dans tous les cas, l'énergie totale émise dans l'espace autour de S est :

$$W = \int_0^{2\pi} d\theta \int_{-\pi/2}^{\pi/2} I(\theta,\varphi) \cos\varphi \, d\varphi \qquad (1)$$

L'énergie totale recueillie au récepteur est :

$$w = \int_0^{2\pi} d\theta \int_{-\pi/2}^{\pi/2} I(\theta,\varphi)\, a(\theta,\varphi) \cos\varphi \, d\varphi \qquad (2)$$

Si nous sommes amenés, par la suite, à partager la sphère trigonométrique en un nombre N de pavés émetteurs fictifs $\Omega_k (1 \leq k \leq N)$ d'aires res-

pectives $\Delta\omega_k$, l'énergie totale émise sera :

$$W = \sum_k \Delta W_k \tag{3}$$

avec :

$$\Delta W_k = \int_{\Omega_k} I(\theta,\varphi) \cos\varphi \, d\theta \, d\varphi \tag{4}$$

et l'énergie totale recueillie :

$$w = \sum_k \Delta w_k \tag{5}$$

avec :

$$\Delta w_k = \int_{\Omega_k} I(\theta,\varphi) \, a(\theta,\varphi) \cos\varphi \, d\theta \, d\varphi \tag{6}$$

4.3.3.- Méthode des sources-images.-

Dans le cas général où un faisceau de rayons issus de la source ponctuelle est réfléchi sur la paroi, les droites portant les rayons réfléchis n'ont pas un point d'intersection commun : les images de la source, en général, n'existent pas. Cependant, il est des cas où ces images existent, soit rigoureusement, soit avec une certaine approximation, et leur considération peut alors être très utile.

On sait que, lorsqu'une paroi réfléchissante est plane, toute source située en avant de cette paroi possède une image, l'image ayant comme position géométrique le symétrique de la source par rapport au plan de la paroi. Quand la paroi est une portion de sphère, les images existent encore, mais seulement de façon approchée, l'approximation étant d'autant meilleure que la distance de la source à la paroi est plus petite relativement au rayon de la sphère. Enfin, si la paroi est une portion de la surface de révolution obtenue avec une conique tournant autour de son grand axe, l'image de la source existe à condition que cette source soit située à un foyer de la conique. En dehors de ce cas, les images n'existent pas.

Un cas particulièrement intéressant est celui où toutes les parois de la salle sont planes (salle polyédrique). La source possède des images par rapport aux plans des parois, et chacune de ces images a à son

tour des images par rapport à tous les plans pour lesquels elle se trouve du côté réfléchissant. De réflexion en réflexion, on peut arriver à un grand nombre de sources, voire une infinité.

Parmi toutes ces sources, il y en a qui voient le récepteur, l'"éclairent" entièrement ou partiellement et qui, par conséquent sont vues, chacune d'au moins un point du récepteur. Quand on dit que le récepteur voit une certaine source image, cela signifie aussi bien que la source réelle S voit une certaine image, entière ou partielle du récepteur. Cette image est vue sous un certain angle solide, qui découpe un certain pavé émetteur sur la sphère trigonométrique attachée à S. On a donc, pour l'ensemble des récepteurs vus de S, c'est-à-dire pour l'ensemble des sources vues du récepteur, un certain partage de la sphère en pavés émetteurs. Chaque pavé correspond à une source; il est caractérisé par l'unité de destinée de tous les rayons qui en émanent. L'effet de mosaïque dont nous avons parlé plus haut est lié au fait que beaucoup de sources images ne voient qu'une partie du récepteur.

Lorsqu'il y a des parois courbes, nous avons énuméré plus haut les cas favorables où, suivant certaines conditions, une source possède, au moins approximativement, une image par rapport à une telle paroi. Cependant ces conditions favorables ne sauraient être vérifiées à la fois par des images distribuées un peu partout dans l'espace et par conséquent, les images successives qu'il était toujours possible de construire lorsque toutes les parois étaient planes n'existent pas ici.

Reconnaissons toutefois que, dans certains cas particuliers présentant un très petit nombre de sources images fortement prépondérantes, la considération de telles sources s'imposera même s'il y a des parois courbes. Mais cette remarque ne contredit pas notre analyse, qui s'adresse à l'ensemble des sources images vues du récepteur. Cet ensemble prétend à la propriété suivante qui n'est pas évidente à priori : tout rayon arrivant au récepteur provient soit directement de la source réelle, soit d'une source image. Ceci exclut l'existence de rayons aberrants.

La méthode exposée ci-après traite de cet ensemble dans une salle polyédrique.

La considération des sources images présente un intérêt théorique et un intérêt pratique. L'intérêt théorique est dans leur répartition

spatiale qui, selon la forme de la salle, présente des symétries ou des périodicités intéressantes. Nous reverrons cela à propos des salles parallélipipédiques. L'intérêt pratique est que la construction des sources images est très facile avec le seul secours de la géométrie élémentaire, ce qui rend aisé leur recensement par un procédé automatique.

On appelle sources images du premier ordre celles qui sont obtenues à partir de la source réelle se réfléchissant une seule fois dans les parois planes de la salle, images du deuxième ordre celles que l'on obtient à partir de celles du premier ordre, et ainsi de suite. L'ordre est une propriété intrinsèque d'une source image : c'est le nombre de symétries qu'il faut effectuer pour la construire.

Le récepteur ne voit pas toutes les sources images. Pour lui, tenant compte de cette restriction, l'ordre d'une source image se définit ainsi : si un point du récepteur voit une source image dans une certaine direction, l'ordre de cette source image est le nombre de réflexions que subit le rayon, issu de la source réelle, qui arrive au récepteur précisément dans cette direction. On peut dire aussi que la direction de départ de ce rayon est celle dans laquelle la source réelle voit une certaine image du récepteur : il y a concordance entre les sources images vues du récepteur et les images du récepteur vues de la source réelle. Selon les besoins du raisonnement, on pourra considérer les phénomènes de l'une ou l'autre façon.

Introduisons une numérotation des sources et soit S_j l'une d'elles : $S_1 = S$ est la source réelle, $j > 1$ correspond aux sources images. De la même façon, soit M_j avec $j > 1$, l'une des images du récepteur vues de S : $M_1 = M$ est le récepteur réel.

La source S voit le récepteur M_j sous l'angle solide Ω_j, de mesure $\Delta\omega_j$. Si le récepteur est vu entièrement, l'angle solide est, dans le cas du récepteur sphérique :

$$\Delta\omega_j = A / 4 R^2_j \qquad (7)$$

où $A=4\pi r^2$ est l'aire du récepteur, r son rayon, et R_j la distance de S_j au centre C du récepteur. Dans le cas du récepteur plan l'angle solide vaut :

$$\Delta\omega_j = Ahj / R^3_j \qquad (8)$$

hj étant la distance de S_j au plan du récepteur.

Les auteurs qui traitent des sources images se placent généralement dans le premier cas, pour tenir compte de la réalité expérimentale : microphone sphérique, ou tête d'un auditeur [3.8].
Les expressions ci-dessus ne sont rigoureuses que si les dimensions du récepteur tendent vers zéro, circonstances qui ont défini notre option b) mais, dans le cas contraire, elles sont une approximation bien suffisante si les dimensions du récepteur sont assez petites vis-à-vis des distances R_j.

Lorsque le récepteur n'est pas vu entièrement de S_j, l'angle solide $\Delta\omega_j$ est une fraction η_j des expressions ci-dessus. Dans tous les cas, l'énergie totale émise W est donnée par l'équation (1). En outre, les expressions (3) et (4) sont valables si l'ensemble des pavés émetteurs Ω_j couvre complétement la sphère trigonométrique.

Le facteur d'affaiblissement $a(\theta,\varphi)$ est le produit des facteurs de réflexion aux points de réflexion rencontrés, entre la source et le récepteur, par le rayon émis dans la direction $D(\theta,\varphi)$. L'énergie totale reçue w est donnée par l'expression (2).
Les expressions (5) et (6) sont certainement valables si l'ensemble des pavés émetteurs recouvre la sphère trigonométrique, mais leur utilisation mérite quelque attention dans le cas de notre option b) . Lorsque l'on diminue les dimensions du récepteur, les $\Delta\omega_j$ diminuent, ce qui entraîne une diminution des ΔW_j. A la limite, tous les ΔW sont nuls, alors que leur somme W n'a pas changé. Que se passe-t-il? la réponse est à chercher dans la répartition des sources images dont l'intensité est, soit nulle, soit très petite à cause de l'éloignement. En effet, lorsqu'on rapetisse le récepteur, certains rayons qui étaient précédemment absorbés par lui, passent à côté de lui et continuent à"tourner" dans la salle jusqu'à ce qu'ils tombent, soit sur une paroi complétement absorbante, soit enfin sur le récepteur. On comprend, pour le premier cas, que les sources images qui conduisent à un facteur a nul et n'ont donc pas à être prises en compte dans (5) et (6) doivent cependant être admises dans (3) et (4).
Pour le second cas, le rapetissement des récepteurs images vus de S démasque progressivement d'autres récepteurs images, de plus en plus lointains, et, par conséquent, de moins en moins visibles, à cause du grand nombre de réflexions. S'il est admis que les sources images pour lesquelles le facteur

a est au-dessous d'un certain seuil n'ont plus à être prises en compte dans (5) et (6), cela n'est pas vrai pour (3) et (4).

En conclusion, les pavés Ω_j utiles pour l'option b) ne couvrent pas toute la sphère trigonométrique et la vérification de W par la relation (3) n'est donc pas possible. Par contre, la relation (5) n'exprime aucune contradiction, puisque les Δw_i qui tendent vers zéro avec les $\Delta\omega_j$ ont une somme w qui tend aussi vers zéro.
Dans le cas de la variante b' , les difficultés ci-dessus ne se présentent pas. La partition de la sphère en pavés Ω_j est la même que dans l'option a) et rien ne change pour les ΔW_j quand on réduit la partie du récepteur considérée comme effectivement réceptrice tandis que le récepteur considéré comme absorbant conserve des dimensions invariables. Pour l'expression (6) les choses se passent à peu près comme dans le cas de l'option b ,mais avec la différence qu'il n'y a pas d'apparition progressive de nouvelles sources. Comme pour l'option b , seules interviennent dans (5) les sources images vues du centre C du récepteur.

Le procédé pratique exposé ci-après présente nécessairement des simplifications par rapport aux développements théoriques précédents. Ces derniers avaient pour but d'éclairer la nature de l'erreur de méthode liée à l'introduction de telles simplifications.
Dans la méthode des sources images, il ne s'agira pas de calculer dans les angles solides Ω_j les intégrales définies (4) et (6) : pour chaque pavé Ω_j on s'intéressera seulement au rayon qui part de S dans la direction $D_j(\theta_j, \varphi_j)$ du centre du récepteur image Mj . Les intégrales précédentes sont alors remplacées par :

$$\Delta W_j = I(\theta_j, \varphi_j)\Delta\omega_j \qquad (9)$$

$$\Delta w_j = I(\theta_j, \varphi_j)\, a(\theta_j, \varphi_j)\Delta\omega_j \qquad (10)$$

Il ne s'agira pas non plus de calculer les facteurs η_j relatifs aux sources images qui n'éclairent pas la totalité du récepteur : ce facteur est à prendre égal à l'unité ou à zéro selon que la source image Sj voit ou non le centre C du récepteur. Cette façon de procéder est celle qui convient à la résolution rigoureuse du problème que nous avons appelé option b .

Soit E la limite du quotient w/A lorsque les dimensions du récepteur, toujours centré au point C, deviennent très petites. Les expressions (9) et (10) sont rigoureuses en tant que limite des expressions (4) et (6), et l'expression (10), utilisée avec (7) ou (8) suivant le cas, permet d'écrire :

$$E = \sum_j \frac{\Delta w_j}{A}$$

ce qui donne, pour le récepteur sphérique :

$$E = \sum_j \frac{1}{4 R^2_j} I(\theta_j, \varphi_j)\, a(\theta_j, \varphi_j)$$

et pour le récepteur plan :

$$E = \sum_j \frac{h_j}{R^3_j} I(\theta_j, \varphi_j)\, a(\theta_j, \varphi_j)$$

Telles sont les expressions qu'il y aura lieu de calculer une fois effectué le recensement des sources images S_j vues du point C. Il est à noter que la numérotation des sources images est ici quelque peu différente de celle qui a été introduite précédemment, mais il ne peut y avoir de confusion.

Un point reste à préciser : comment savoir si une source image est visible de C? Nous désignerons par $S_{(m)}$ la source, d'ordre m à essayer. Construisons le point Q_m, intersection de la droite $CS_{(m)}$ avec le plan π_m qui a servi à la dernière réflexion. Si le point Q_m n'appartient pas à la portion de plan π_m formant la paroi de la salle, la source $S_{(m)}$ n'est pas visible de C. Si Q_m appartient à la paroi, on construit alors l'intersection Q_{m-1} de la droite $Q_s S_{(m-1)}$ avec la plan π_{m-1} qui a servi à l'avant-dernière réflexion, $S_{(m-1)}$ étant la source image qui, par réflexion sur π_m a donné $S_{(m)}$. Si Q_{m-1} n'appartient pas à la paroi, la source $S_{(m)}$ est invisible de C. Dans le cas contraire, on remonte de proche en proche à des points d'intersection qui doivent tous appartenir aux parois pour que la source image $S_{(m)}$ soit réputée visible de C.

Il faut bien observer que, si une source image d'ordre m n'est pas vue de C, elle ne doit pas pour autant être abandonnée car il reste possible que, parmi les sources images d'ordre supérieur qu'elle engendre, certaines soient visibles.

Dans le cas où le problème à résoudre est ce que nous avons appelé la variante b' , une vérification supplémentaire est à faire à chaque étape du procédé ci-dessus : les segments de droite $Q_m S_{(m-1)}$, $Q_{m-1} S_{(m-2)}$, ... , $Q_1 S$ ne doivent pas rencontrer le récepteur, celui-ci étant supposé absorbant et de dimensions finies.
La connaissance de la direction initiale D_j du rayon ainsi reconstitué ne demande aucune détermination supplémentaire puisque c'est la direction du segment SQ_1. Le cas d'une source anisotrope ne présente donc pas de difficulté particulière. Toutefois, dans nos essais numériques, nous nous sommes bornés au cas de la source isotrope.

a) - Cas de la salle parallélipipédique.-

La simplification des calculs, alliée à la simplification des phénomènes eux-mêmes, donne un intérêt particulier au cas d'une salle rectangulaire.

On vérifie aisément que, lorsque la source est à l'intérieur d'un dièdre aigu , les deux images d'ordre deux, c'est-à-dire obtenues après deux réflexions, sont vues simultanément de toute une région à l'intérieur du dièdre, tandis que dans les autres régions on n'en voit qu'une sur les deux. Lorsque le dièdre est obtus, la région où on les voit toutes deux n'existe plus; il y a, par contre, une région où aucune des deux n'est vue. Quand le dièdre est droit, les deux régions particulières ci-dessus disparaissent, on ne voit jamais qu'une seule source image d'ordre deux et il n'y a pas de discontinuités dans les éclairements.
Les dièdres droits amènent donc d'intéressantes simplifications dans le recensement des sources images vues du récepteur.

Une salle dont tous les dièdres sont droits, c'est-à-dire un parallélipipède rectangle, est caractérisée par la continuité spatiale des éclairements. Les images de la salle par rapport à ses parois remplissent l'espace en formant un réseau dont la maille est la salle elle-même. Les sources images sont situées suivant des plans réticulaires; il y en a une par maille du réseau. Leurs coordonnées cartésiennes sont particulièrement aisées à calculer.

Certains auteurs ont été encouragés par la simplicité que présente l'étude d'une telle salle. Ils l'ont approfondie en faisant diverses hypothèses sur la répartition du facteur de réflexion : cas où tou-

tes les faces sont parfaitement réfléchissantes sauf une parfaitement absorbante, cas de deux faces absorbantes opposées ou adjacentes, etc... Selon le cas, le réseau des sources images se trouve restreint à deux plans parallèles, ou à un seul plan, ou à des domaines encore moins étendus.

b) - Cas d'une salle quelconque.-

Une salle peut avoir la forme d'un polyèdre irrégulier quelconque; c'est le cas auquel s'adapte parfaitement la méthode que nous avons exposée. Pour une salle comportant une ou plusieurs parois courbes, nous avons souligné l'impossibilité d'y considérer un ensemble de sources images tel qu'il est utilisé dans notre méthode.

On peut penser réduire aussi l'impossibilité; discrétiser l'ensemble des parois en autant de portions planes qu'il faudra pour respecter ces parois dans une marge d'approximation donnée et être ainsi ramené au cas d'une salle polyédrique.

C'est ici qu'il y a lieu de prêter attention à un aspect pratique de l'application de la méthode; il s'agit de la prolifération des sources images en fonction de l'ordre. Le recensement des sources images vues du récepteur comporte, pour chacune d'elles, sa construction à partir d'une source d'ordre inférieur, puis la vérification qu'elle est effectivement vue du récepteur. Or, si l'on considère que le nombre total de sources images retenues jusqu'à l'ordre m croît tout au plus comme m tandis que le nombre total de sources à essayer peut, très grossièrement, être estimé à $(p/2)^m$, p étant le nombre de plans constituant les parois, on se rend compte du déchet croissant rencontré quand l'ordre m croît.

Il n'est donc pas possible, pratiquement, de pousser bien loin la discrétisation des parois. Il apparaît d'autre part, et ceci est également vrai pour une salle réellement polyédrique, qu'il faut être modeste dans le choix de l'ordre maximal : les applications numériques ont montré que, au-delà d'une certaine valeur de m, la légère amélioration des résultats apportée par l'accroissement de m d'une unité devenait très onéreuse en temps de calcul. On voit ici l'intérêt pratique de tirer parti de l'existence éventuelle des dièdres droits, dont nous avons rappelé les propriétés au paragraphe précédent.

Cependant, si le recensement des sources images individuelles par la méthode que nous avons exposée n'est pas possible au-delà d'un certain ordre m, donc pratiquement au-delà d'une certaine distance R du récepteur, il est toujours possible de considérer le nuage des sources images non calculées et de faire des hypothèses sur leur distribution. Ces hypothèses permettront, une fois introduit l'aspect temporel, de prolonger, au moins approximativement, les résultats obtenus avec les sources recensées individuellement, en particulier en ce qui concerne l'extinction lointaine du son dans une salle.

4.3.4.- Méthode des rayons discrets.-

La méthode précédente revenait à déterminer à priori les directions de départ des rayons qui, après certaines réflexions, allaient passer exactement par le centre du récepteur. Nous ne ferons plus maintenant une telle détermination préalable; nous ferons partir des rayons de S dans des directions arbitraires et nous suivrons chacun d'eux au cours de ses réflexions successives jusqu'à son arrivée sur le récepteur ou à sa complète absorption. Le principe de la méthode des rayons discrets est de distribuer toute l'énergie émise par la source S à un nombre fini de rayons partant dans des directions toutes différentes. L'énergie transportée par chaque rayon se propage suivant une certaine demi-droite jusqu'à la rencontre soit du récepteur, soit d'une paroi de la salle. Le premier cas a lieu lorsque le point d'intersection du rayon avec la paroi appartient à la portion de paroi définie comme récepteur, ou, si le récepteur est sphérique, lorsque la distance du rayon au centre de la sphère est inférieure au rayon de la sphère. L'énergie incidente est alors mise au compte de l'énergie reçue et l'on passe au rayon suivant. Dans le second cas, s'il ne s'agit pas d'une paroi complètement absorbante, il y a lieu de construire le rayon réfléchi; cela est possible puisque par hypothèse, on connaît la direction de la normale en tout point de la paroi, donc au point de réflexion. Connaissant également le facteur de réflexion, fonction ou non de l'angle d'incidence, on en déduit l'énergie du rayon réfléchi.

Ainsi, la position et l'énergie de chaque rayon sont exactement connus après chaque réflexion, et cela jusqu'à ce qu'il rencontre le récepteur, apportant alors sa contribution à l'énergie reçue. S'il arrive

qu'un rayon, avant de rencontrer le récepteur ou une paroi complétement absorbante, ait cependant subi un nombre de réflexions tel qu'il ne transporte plus qu'une énergie inférieure à un seuil préalablement fixé, la poursuite de ce rayon est abandonnée. A ce propos, une comparaison peut être faite avec la méthode des sources images : le seuil en question peut correspondre à un nombre de réflexions supérieur à celui qu'autoriserait pratiquement la méthode des sources images, si bien que la méthode des rayons discrets offre, bien mieux que l'autre méthode, la possibilité d'étudier la phase d'extinction lointaine.

Observons que le passage exact d'un rayon par le point C ne peut être qu'exceptionnel et que la notion de limite introduite dans la définition de notre option b n'a ici aucune signification. La méthode des rayons discrets exposée ici ne peut donc viser qu'à résoudre l'option a .

Numériquement, l'application de la méthode des rayons discrets consiste, une fois définie la distribution de l'énergie, au départ, entre les différents rayons, dans la résolution réitérée des quatre problèmes de géométrie analytique suivants :

- a) - équation d'une demi-droite issue d'un point donné et ayant une direction donnée,
- b) - intersection d'une demi-droite et d'une surface donnée par son équation (ou un système d'équations),
- c) - appartenance d'un point à une portion de surface définie par une inéquation ou un système d'équations et, pour le cas du récepteur sphérique, distance d'un point à une demi-droite,
- d) - direction du rayon réfléchi, symétrique du rayon incident par rapport à la normale à la surface au point d'intersection calculé. Il s'y ajoute, éventuellement, le calcul du facteur de réflexion d'après l'angle d'incidence.

La méthode des rayons discrets peut s'appliquer à n'importe quelle forme de salle. Il suffit de savoir mettre en équations l'ensemble des parois de la salle et de savoir résoudre le problème b ci-dessus avec ces équations.

a) - Distribution de l'énergie aux rayons discrets.-

Nous avons posé en principe de la méthode que nous choisirions les directions des rayons au départ de S; il faut maintenant préciser comment sera fait ce choix. Définissons une partition de la sphère trigonométrique en un nombre arbitraire N de pavés émetteurs. Nous choisirons sur chaque pavé Ω_k $(1 \leq k \leq N)$, d'aire $\Delta\omega_k$, le point directeur $D_k(\theta_k, \varphi_k)$ d'un rayon auquel nous attribuerons au départ toute l'énergie ΔW_k du pavé, comme par la relation (4). Ce rayon étant suivi tout au long de ses réflexions successives, on calcule son facteur d'affaiblissement $a(\theta_k, \varphi_k)$ et il est censé apporter au récepteur une énergie :

$$\Delta w_k = a(\theta_k, \varphi_k)\Delta W_k \qquad (11)$$

Cette expression est équivalente à l'expression (6) dans le cas où tous les rayons issus du pavé Ω_k ont une même "destinée".
Cela est possible, mais il est possible, aussi, que la fonction $a(\theta,\varphi)$ ait sur Ω_k plusieurs valeurs très différentes car, comme nous l'avons dit, cette fonction présente de nombreuses discontinuités.
Pour calculer par l'expression (5) l'énergie totale reçue w , nous ne chercherons jamais à déterminer les Δw_k par l'expression (6), mais nous utiliserons les valeurs données par (11). On sait qu'en opérant de cette façon simplifiée on obtient une valeur inexacte de l'énergie totale reçue. La question se pose alors de savoir estimer l'erreur que l'on commet en procédant ainsi et éventuellement, de savoir comment la réduire.

Plusieurs manières de définir les directions de départ D_k des rayons discrets et leurs énergies ΔW_k sont possibles. On pourrait adopter comme point directeur le milieu du pavé Ω_k ; l'inconvénient de ce procédé est la difficulté d'estimer l'erreur. Nous préférons choisir chaque point D_k au hasard sur le pavé Ω_k correspondant, suivant une loi de probabilité uniforme. Si l'on recommence le calcul plusieurs fois, les résultats des tirages au hasard sont différents et la dispersion des valeurs du résultat final w renseigne sur la dispersion de l'erreur.

Les N quantités $a_k = a(\theta_k, \varphi_k)$ servant à calculer les Δw_k par l'expression (11) sont alors des variables aléatoires; il en est de même des Δw_k. Quant aux ΔW_k , nous avons deux façons de les définir : utiliser leurs valeurs exactes connues par (4), où le calcul de l'intégrale ne doit pas présenter de difficultés, ou bien se servir de l'expression (9) utilisée dans la méthode des sources images, mais en prenant comme valeur

de l'intensité d'émission celle $I_k = I(\theta_k, \varphi_k)$ qui correspond à chaque point D_k résultant du tirage au hasard. Ce dernier procédé a peut-être l'inconvénient d'accroître la dispersion des Δw_k en remplaçant des I_k certains par des variables aléatoires; nous le trouvons cependant avantageux car, outre que la dispersion supplémentaire introduite est généralement faible vis-à-vis de celle qui est due aux facteurs a_k, les valeurs probables des variables aléatoires définies de cette façon sont égales aux valeurs exactes données par (6). Nous calculerons donc w par l'estimateur :

$$w' = \sum_k I(\theta_k, \varphi_k)\, a(\theta_k, \varphi_k) \Delta\omega_k$$

Nous connaissons la moyenne de la variable aléatoire w', c'est w.
Nous voudrions aussi connaître son écart-type $\sigma(w')$ ou sa variance $v(w')$
$v(w') = \sigma^2(w')$ pour caractériser la dispersion de w' autour de w.

Pour le calcul des dispersions, nous ne tiendrons pas compte de celle qui est due aux facteurs I_k. On a :

$$\sigma(\Delta w'k) = I_k \Delta\omega_k \sigma(a_k)$$

et comme les a_k sont des variables aléatoires mutuellement indépendantes, la variance de la somme est la somme des variances des termes :

$$v(w') = \sum_k (Ik\, \Delta\omega k)^2 \; v(a_k)$$

Nous désirons minimiser cette expression : il y a lieu de faire ici une hypothèse raisonnable sur la façon dont les $v(a_k)$ dépendent des dimensions et de la forme des pavés. Posons que l'écart-type $\sigma(a_k)$ est proportionnel à une certaine puissance ρ_k^{α} $(\alpha > 0)$ du rayon de giration ρ_k du pavé Ω_k, le coefficient de proportionnalité étant le même pour tous les pavés.
Une conséquence immédiate de cette hypothèse est la suivante :

1) - Le partage de la sphère trigonométrique en N pavés émetteurs fictifs doit être fait de telle sorte que chacun d'eux, pour l'aire $\Delta\omega k$ qui lui est attribuée, possède, par rapport à son centre de gravité, un moment d'inertie aussi petit que possible.
Pour l'attribution des $\Delta\omega_k$, il y a lieu de tenir compte de la distribution de l'intensité de la source. La question qui se pose est celle-ci :

dans les directions (régions de la sphère) où l'intensité I est plus grande, faut-il réduire les pavés pour qu'ils soient tous porteurs d'une même énergie, ou faut-il que cette énergie soit une fonction croissante de I, ou décroissante?

Nous pouvons écrire la variance $v(w')$ sous la forme :

$$v(w') = \sum_{k=1}^{N-1} (I_k \Delta\omega_k)^2 \, v(a_k) + I_N^2 \, v(a_N) \left(4\pi - \sum_{k=1}^{N-1} \Delta\omega_k\right)^2$$

le dernier pavé $\Delta\omega_N$ étant déduit des $N-1$ précédents :

$$\Delta\omega_N = 4\pi - \sum_{k=1}^{N-1} \Delta\omega_k$$

Par contre, les $\Delta\omega_k$ sont des variables indépendantes pour $1 \leqslant k \leqslant N-1$. Ecrivons alors que :

$$\frac{dv(w')}{d(\Delta\omega_k)} = 0 \quad \text{pour} \quad 1 \leqslant k \leqslant N-1$$

On obtient : $2(I_k^2 \Delta\omega_k) v(a_k) - 2I_N^2 \, v(a_N) \left(4\pi - \sum_{k=1}^{N-1} \Delta\omega_k\right) = 0$

soit $N-1$ égalités :

$$\begin{aligned} I_1^2 \Delta\omega_1 v(a_1) &= I_N^2 \Delta\omega_N \, v(a_N) \\ &\vdots \\ I_{N-1}^2 \Delta\omega_{N-1} v(a_{N-1}) &= I_N^2 \Delta\omega_N \, v(a_N) \end{aligned}$$

comme $v(a_k) \simeq \rho_k^{2\alpha}$ et $\Delta\omega_k \simeq \rho_k^2$, on aboutit au résultat suivant :

$$I_k \, \rho^{\alpha+1} = \text{cte}$$

ce qui peut s'écrire également sous la forme :

$$\Delta W_k \cdot I_k^{\frac{1-\alpha}{1+\alpha}} = \text{cte} \text{ , étant donné que } \Delta W_k = I_k \cdot \Delta\omega_k \text{ .}$$

Revenons à l'expression de la variance $v(w')$ et appelons C la constante

$$I_k^2 \Delta\omega_k \; v\, a_k = I_k^2 \Delta\omega_k^{(1+\alpha)}$$

Il vient

$$v(w') = \sum (I^2_k \Delta\omega_k \; v\, a_k)\Delta\omega_k = C\sum\omega_k = 4\pi C.$$

D'autre part, la somme :

$\sum I^{2/(1+\alpha)} \Delta\omega_k$ peut être assimilée à une intégrale définie égale à une certaine constante $A = N.C$

On obtient par conséquent :

$$v(w') = 4\pi\,(A/N)^{1+\alpha}$$

On peut alors énoncer le résultat suivant :

2) - L'aire des pavés doit être prise inversement proportionnelle à $I^{[2/(1+\alpha)]}$, c'est dire que les produits $I^{\frac{1-\alpha}{1+\alpha}} \Delta W$ doivent être les mêmes pour tous les pavés. La variance $v(w')$ est alors inversement proportionnelle à $N^{1+\alpha}$.

Le petit nombre de nos essais numériques ne nous a pas permis de donner la préférence à une valeur de l'exposant α qui, d'ailleurs, ne se prête pas à une détermination théorique évidente.
Le choix $\alpha = 1/2$ traduirait une hypothèse selon laquelle la variance $v(a_k)$ serait proportionnelle à ρ_k, le choix $\alpha = 1$ supposerait la proportionnalité de l'écart type $\sigma(a_k)$ à ρ_k.
Ce dernier choix donne une forme particulièrement simple à la deuxième règle de partage : tous les ΔW_k doivent être égaux, et l'écart-type $\sigma(w')$ est alors proportionnel à $1/N$.

Le respect rigoureux des deux règles précédentes n'est pas essentiel; d'ailleurs, il est pratiquement impossible. Cependant, en tenant compte de ces règles, on pourra réduire autant que possible les incertitudes inhérentes à la méthode des rayons discrets.

Ajoutons deux remarques.

- La première concerne conjointement l'étendue du récepteur et le choix du nombre N. Le besoin de réduire autant que possible l'écart-type $\sigma(w')$ est déjà une bonne raison de choisir N assez grand. En outre, il faut observer que le nombre des rayons tombant sur le récepteur n'est qu'une petite fraction du nombre de rayons émis et que, si l'on veut que les

résultats signifient quelque chose, ces rayons-là ne doivent pas se réduire à quelques unités. Il faut donc, en premier lieu que N soit assez grand (cf. nos applications numériques), en second lieu que soit assez grande l'étendue de ce qu'il est convenu d'appeler "récepteur", ou paroi réceptrice, condition qui contraste radicalement avec celle que nous avons dû poser à propos de la méthode des sources images.

- Notre seconde remarque concerne le calcul du rapport de l'énergie reçue à l'énergie émise. Le calcul des ΔW_k à partir des I_k plutôt que par l'expression (4) a pour effet de remplacer la valeur certaine W de l'énergie émise par la source S par une variable aléatoire W'.
On aura intérêt à rapporter l'énergie reçue à W' plutôt qu'à W, même si cette dernière quantité est exactement connue. En effet, grâce à la corrélation qui existe entre W' et w', la variance de la variable aléatoire w'/W' est un peu inférieure à celle de la variable aléatoire w'/W ; la première est donc préférable. Cette dernière remarque, comme certains autres détails de ce paragraphe, n'est évidemment d'aucun intérêt dans le cas où l'intensité I est uniforme.

b) - Cas de la source isotrope.-

Quand l'intensité I est la même dans toutes les directions, l'énergie totale émise est $W = 4\pi I$. La deuxième règle énoncée au paragraphe précédent nous a conduit à diviser la sphère trigonométrique en N pavés équivalents, chacun ayant donc l'aire $4\pi/N$ et l'énergie W/N.

Le respect rigoureux de la première règle de partage conduit à diviser la sphère d'une certaine façon en pavés pentagonaux et hexagonaux; cette solution est bien trop compliquée numériquement pour l'usage que nous avons à en faire. La solution ci-dessous ne conduit pas très loin du minimum cherché de $\sigma(w')$ et surtout elle est relativement facile à appliquer.

Dans le système de coordonnées sphériques utilisé précédemment, nous découperons chaque hémisphère en zones, dont les largeurs sont à déterminer, et chaque zone en pavés égaux par des arcs de méridiens équidistants. Les pavés sont donc des rectangles sphériques, à l'exception de quelques-uns situés aux deux pôles. Pour minimiser les rayons de giration, nous choisirons les dimensions de ces rectangles pour qu'ils soient aussi peu différents que possible d'un carré, ce qui revient, pour une aire donnée, à minimiser le périmètre. Le côté du carré équivalent à chaque pavé est $\sqrt{4\pi/N}$.

Nous supposerons que N est pair et nous poserons :

$$N' = N/2$$

Soit Φ_{i-1} et Φ_i les latitudes des parallèles qui limitent la i ième zone à partir de l'équateur, q_i le nombre de pavés de cette zone.Les latitudes Φ_i seront déterminées par récurrence, la première étant $\Phi_0 = 0$. Soit n le nombre total de pavés situés entre l'équateur et la latitude Φ_{i-1} et supposés déjà construits :

$$n = \sum_{j=1}^{i-1} q_j$$

Tous les pavés ayant la même aire, nous pouvons écrire, d'après la géométrie élémentaire :

$$\sin \Phi_{i-1} = n/N'$$

La latitude Φ_{i-1} est donc connue. Nous voulons partager la longueur $2\pi \cos \Phi_{i-1}$ du parallèle en q_i parties égales, l'entier q_i étant tel que le périmètre des pavés dela i ième zone soit minimal.
Une solution est fournie par la formule que nous proposons : on prendra pour q_i la partie entière du nombre réel donné par l'expression :

$$\sqrt{\frac{2\pi}{N'}\ (N'^2 - n^2)} - \frac{3n}{N'} + \frac{1}{2}$$

Une fois effectué le découpage de la sphère trigonométrique en N pavés, il s'agit de tirer au hasard, sur chaque rectangle sphérique Ω_k, le point directeur D_k . Sur chaque pavé de la i ième zone, on tirera au hasard $\sin \varphi_k$ sachant que sa probabilité est uniforme entre $\sin \Phi_{i-1}$ et $\sin \Phi_i$, c'est-à-dire entre n/N' et $(n + q_i)/N'$. Sur le j ième pavé de cette zone, on tirera au hasard θ_k sachant que sa probabilité est uniforme entre $2\pi (j-1)/q_i$ et $2\pi j/q_i$.

4.5.4.- Introduction de l'aspect temporel.-

Supposons maintenant que l'énergie est émise par la source exactement à l'instant $t = 0$. Nous allons étudier comment s'échelonne dans le temps l'énergie captée au récepteur. La vitesse c de propagation

du son dans l'air n'aura qu'un rôle de facteur d'échelle pour raccorder au temps t les longueurs R des rayons sonores, considérés jusqu'ici sous leur aspect purement géométrique.

Dans la méthode des sources images, ces longueurs sont définies par les distances des différentes sources au centre C du récepteur. Ces distances peuvent être classées par ordre de grandeur croissante; l'indice j sera maintenant utilisé pour désigner le numéro d'ordre de chaque source dans ce classement. Les temps correspondants :

$$tj = Rj/C$$

sont évidemment classés dans le même ordre, l'énergie à l'arrivée correspondant à la source Sj est Δwj .

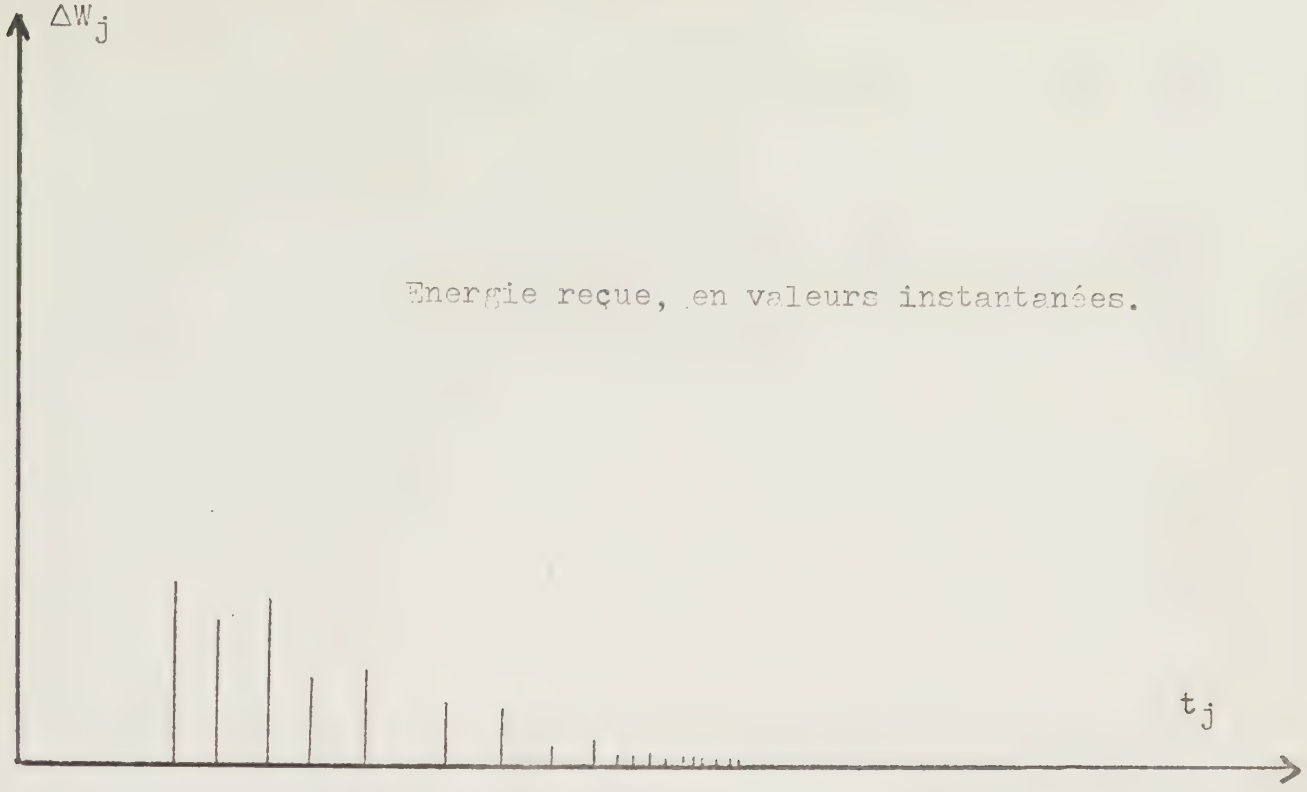
Le diagramme temporel de l'énergie reçue peut être tracé, soit en valeurs instantanées dans le plan des coordonnées (tj , wj), c'est l'"échogramme" (Fig. 4.1), soit en valeurs cumulées dans le plan (t, w) (Fig. 4.2):

$$w(t) = \sum_{j' \leqslant j} \Delta wj' \qquad \text{pour} \quad tj \leqslant t < tj + 1$$

De la même façon, les différents rayons discrets considérés dans la méthode correspondante peuvent être classés par ordre de longueur croissante et l'indice k sera utilisé pour désigner le numéro d'ordre de chacun dans ce classement. Soit $\Delta w'_k$ l'énergie à l'arrivée du rayon discret classé k ième et t'_k son temps d'arrivée. De façon similaire à ce qui a été fait pour l'autre méthode, on peut tracer dans le plan des coordonnées $(t'_k , \Delta w'_k)$ et (t, w') les diagrammes temporels de l'énergie reçue, en valeurs instantanées (Fig. 4.3) et en valeurs cumulées (Fig. 4.4)

Les diagrammes $\Delta w_j (tj)$ et $\Delta w'_k(t'_k)$ méritent d'être attentivement comparés. Supposons que le diagramme $\Delta wj (tj)$ n'est pas celui que l'on a obtenu par l'expression simplifiée (10) de notre méthode des sources images en vue de résoudre l'option b , mais que l'on a voulu résoudre l'option a en utilisant les sources images et l'expression exacte (6). Retournons à la notion de"destinée" : chaque destinée possible correspond à une source Sj vue au moins d'un point du récepteur, ainsi qu'à une image Mj entière ou partielle du récepteur, vue de S sous l'angle solide Ωj . De plus, les longueurs des rayons ayant une même destinée sont comprises dans un intervalle relativement étroit; il en est de même des temps correspondants. Les angles solides Ωj des images les plus proches sont,

Méthode des images, courbes schématiques.



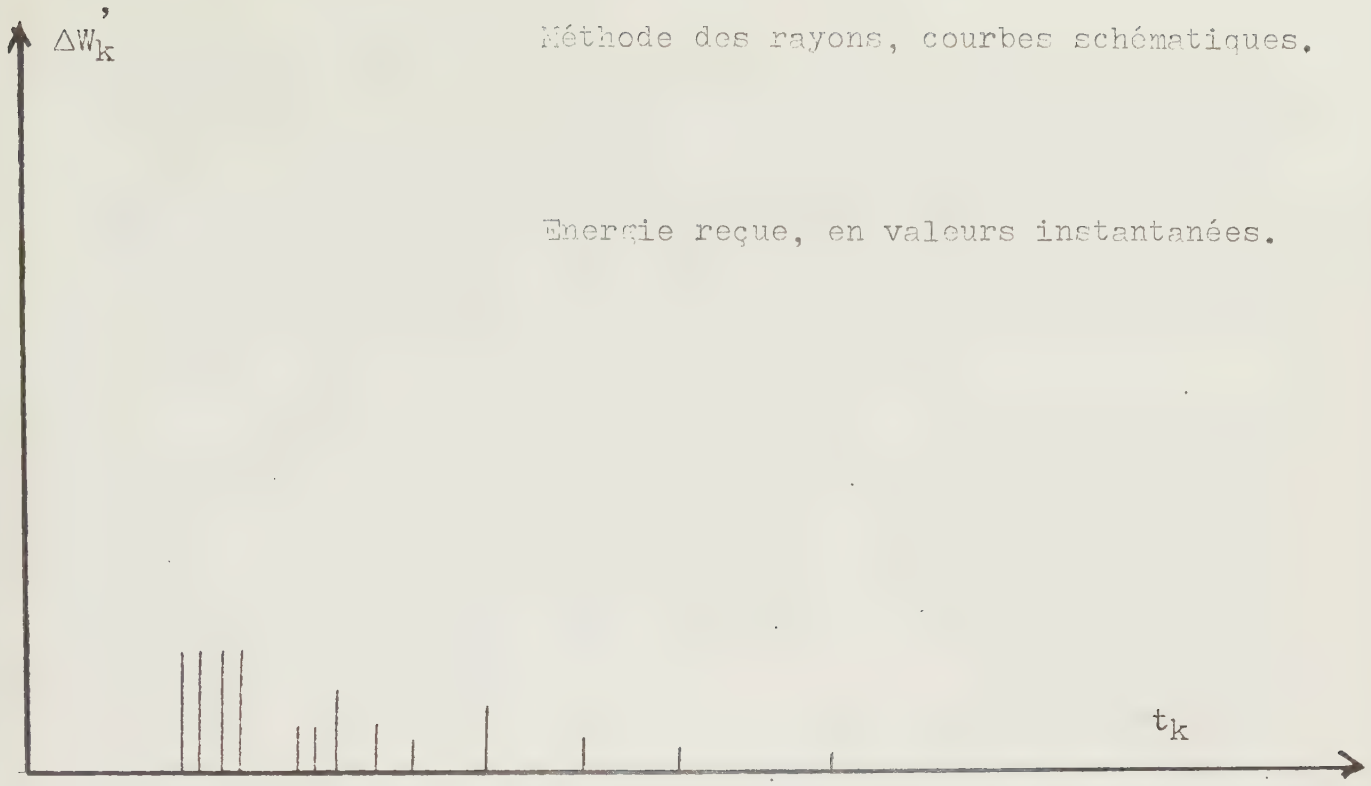
Energie reçue, en valeurs instantanées.

Figure 4.1

Energie reçue, en valeurs cumulées.

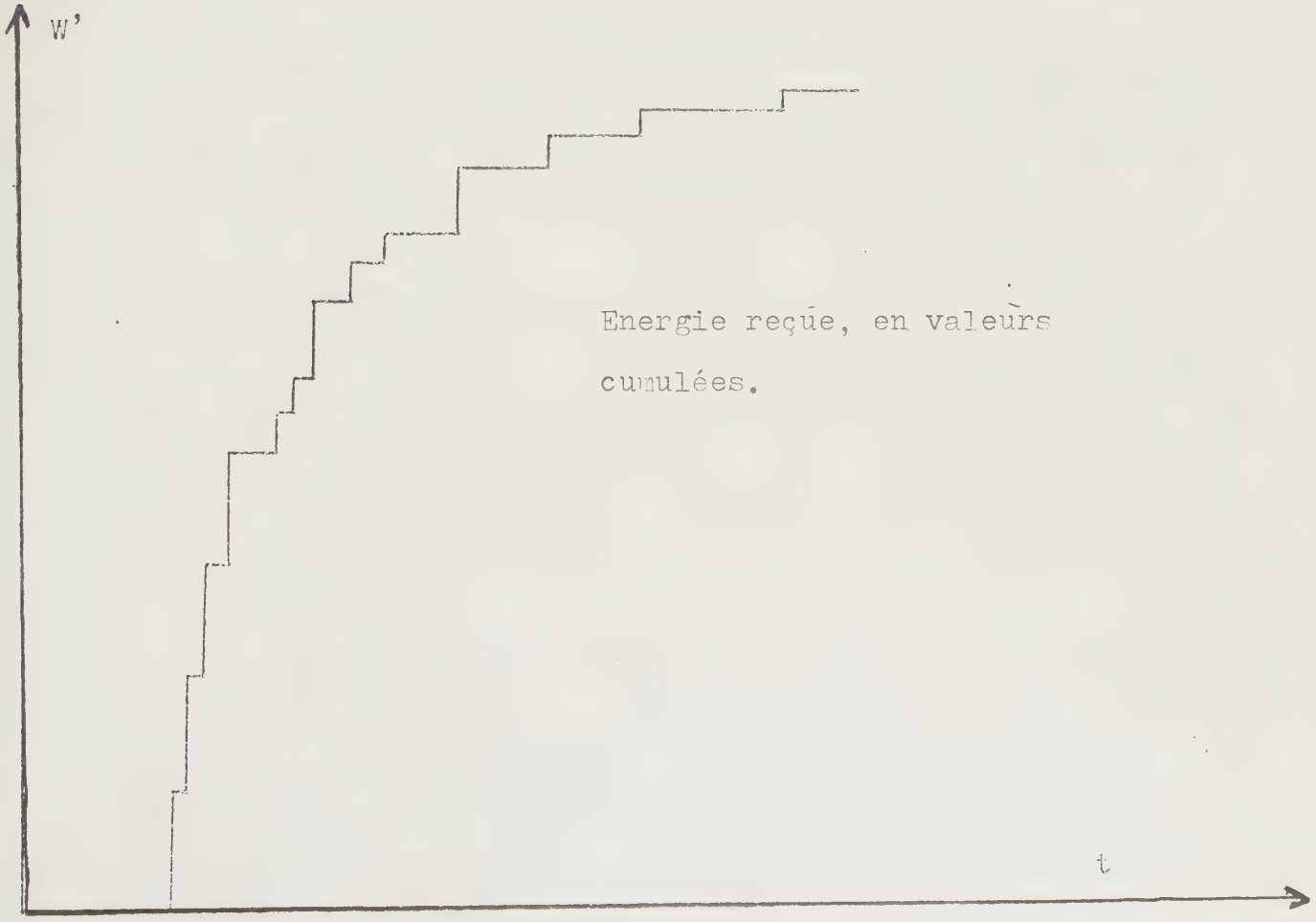
Figure 4.2

Méthode des rayons, courbes schématiques.



Energie reçue, en valeurs instantanées.

Figure 4.3



Energie reçue, en valeurs cumulées.

Figure 4.4

en moyenne, les plus grands et ils peuvent alors contenir un plus grand nombre de pavés Ω'_k utilisés dans la méthode des rayons discrets. Au contraire, les images plus lointaines sont vues sous des angles solides Ωj plus petits et un seul pavé Ω'_k peut contenir plusieurs de ces Ωj. Comme il n'y a jamais qu'un seul rayon discret par pavé Ω'_k, la proportion d'images lointaines atteintes chacune par un rayon discret est d'autant plus faible que la distance R est plus grande.

Revenons maintenant aux Δwj de la Fig. 4.1. Ils diffèrent de ceux du raisonnement précédent en ce sens que les sources images Mj partielles sont traitées comme vues en totalité ou pas du tout selon que l'image du point C est vue ou non. Mais la différence entre les deux distributions d'images, bien que réelle, n'est pas fondamentale et la conclusion, que l'on vérifie en comparant les figures 4.1 et 4.3, résulte du raisonnement précédent : aux petites valeurs de t, les $\Delta w'_k$ données par la méthode des rayons sont plus petits que les Δwj donnés par la méthode des images, mais les t'_k sont plus serrées que les tj ; par contre, lorsque t croît, les Δwj sont de plus en plus petits et les tj de plus en plus serrés, tandis que les t'_k s'espacent de plus en plus.

Quant aux diagrammes w(t) et w'(t) des énergies cumulées, on peut s'attendre à ce qu'ils soient assez voisins l'un de l'autre mais pas nécessairement identiques puisque l'un concerne l'éclairement énergétique au point C, l'autre l'éclairement énergétique moyen du récepteur.

a) Champ d'application et erreurs des deux méthodes.-

Dans les paragraphes suivants, nous désignerons les diagrammes temporels uniformément par $\Delta w_k(t_k)$ et w(t) sans égard à la méthode utilisée pour les obtenir. Nous récapitulerons les diverses observations faites jusqu'ici concernant les possibilités comparées des deux méthodes, indépendamment, pour l'instant, de la notion d'efficacité. Nous admettrons, toutefois, la nécessité de limiter à une certaine valeur m l'ordre des sources images.

La méthode des rayons est applicable à toute forme de salle. Elle résoud seulement l'option a : réponse moyenne de la salle sur l'étendue globale du récepteur. L'erreur qu'elle présente est une erreur d'échantillonage. Cette erreur peut être réduite par l'accroissement du nombre des rayons et l'étendue du récepteur. Enfin, la distribution de cette erreur peut être estimée, et les moyennes améliorées par la réitération des calculs,

utilisant des tirages au sort différents.

La méthode des images est applicable seulement aux salles polyédriques. Elle résoud seulement l'option b : réponse de la salle au centre du récepteur. Au sujet de l'erreur, pour une valeur donnée m de l'ordre maximal des sources images, on peut trouver deux valeurs R_1 et R_2 de la distance R possédant les propriétés suivantes : pour $t < R_1/c$ la solution est rigoureuse, l'erreur de la méthode est nulle. Pour $t > R_2/c$ la méthode ne recense plus aucune source image, elle est totalement en défaut. Entre R_1 et R_2 la signification des résultats se dégrade progressivement. Sans doute peut-on pallier la défaillance de la méthode en faisant, pour les grandes valeurs de R des hypothèses complémentaires sur la distribution des sources images, mais cela n 'est plus du ressort de la méthode exposée.

En d'autres termes, l'étude des échos proches dans une salle polyédrique est rigoureuse par la méthode des images. Dans une salle quelconque, cette étude est possible par la méthode des rayons et l'erreur d'échantillonnage peut être estimée en distribution; elle est aussi possible par la méthode des images, mais l'erreur de discrétisation des parois est difficile à évaluer. L'étude de la phase d'extinction lointaine n'est possible que par la méthode des rayons, les sources images ne restant utiles que par les hypothèses que l'on peut faire sur leur distribution. Aux temps intermédiaires, la méthode des rayons montre une précision à peu près uniforme tandis que la méthode des images voit son erreur naître et croître rapidement à partir d'une certaine valeur de t, fonction de l'ordre maximum m imposé aux sources images.

A ces raisons théoriques, nous adjoindrons des raisons pratiques d'efficacité, importantes en particulier pour fixer d'une part le nombre N de rayons discrets, d'autre part et surtout l'ordre maximal des sources images.

b) Réponse à une impulsion rectangulaire.-

Les fonctions $\Delta w_k(t_k)$ et $w(t)$ seront maintenant considérées comme rapportées à l'énergie émise, ce qui revient à prendre $W = 1$; la fonction $\Delta w_k(t_k)$, distribution temporelle de l'énergie, encore appelée "réponse impulsionnelle" ou "échogramme" est la réponse à une émission sonore instantanée, d'énergie unité. La fonction $w(t)$ est la réponse à un échelon de

Heaviside; la puissance unité est émise par la source S, maintenue indéfiniment à partir de l'instant zéro. La grandeur w est alors une puissance.

Une impulsion rectangulaire émise par S entre les instants t=0 et $t=\tau$ peut être considérée comme la somme algébrique de deux échelons, l'un positif, émis à partir de t=0, l'autre négatif, émis à partir de $t=\tau$. Nous ne faisons aucune hypothèse sur la structure harmonique de l'impulsion, mais nous supposerons que sa puissance constante est égale à l'unité. La puissance arrivant au récepteur à l'instant t est :

$$P(t) = w(t) - w(t - \tau)$$

Observons qu'il suffit de déterminer la fonction w(t) une fois pour toutes : des impulsions rectangulaires de durées différentes se traitent en changeant simplement la valeur de τ dans l'expression ci-dessus.

c) Réponse à une excitation quelconque.-

Supposons que la source S, tout en conservant la même loi de répartition directionnelle, émette un signal dont la puissance est s(t). La réponse de la salle est alors le produit de convolution :

$$P(t) = (S * \Delta w)(t) = \sum_k \Delta w_k \, s(t - t_k) \qquad (12)$$

Pratiquement, il n'est pas nécessaire d'individualiser tous les $\Delta w_k(t_k)$; on peut les grouper par tranches de durée constante, par exemple 1ms ou 10ms.

Ici encore, observons que la détermination de la fonction $\Delta w(t)$ n'est nécessaire qu'une seule fois pour traiter tous les signaux s(t) que l'on voudra.

d) Intégration auditive d'un signal.-

Divers travaux expérimentaux sur l'audition ont montré que, lorsqu'un stimulus acoustique instantané frappait l'oreille, la réponse de celle-ci n'était pas un phénomène instantané : la sensation auditive est distribuée dans le temps suivant une certaine loi, $s_0(t)$ [4.29] .

Adoptons, par exemple la loi :

$$S_0(t) = t^2 e^{-\alpha t}$$

où le paramètre α dépend sans doute de l'auditeur, mais surtout de certains attributs du son entendu, entre autres sa structure harmonique et son niveau sonore. La Fig. 4.5 montre les variations de la fonction exprimée ci-dessus.

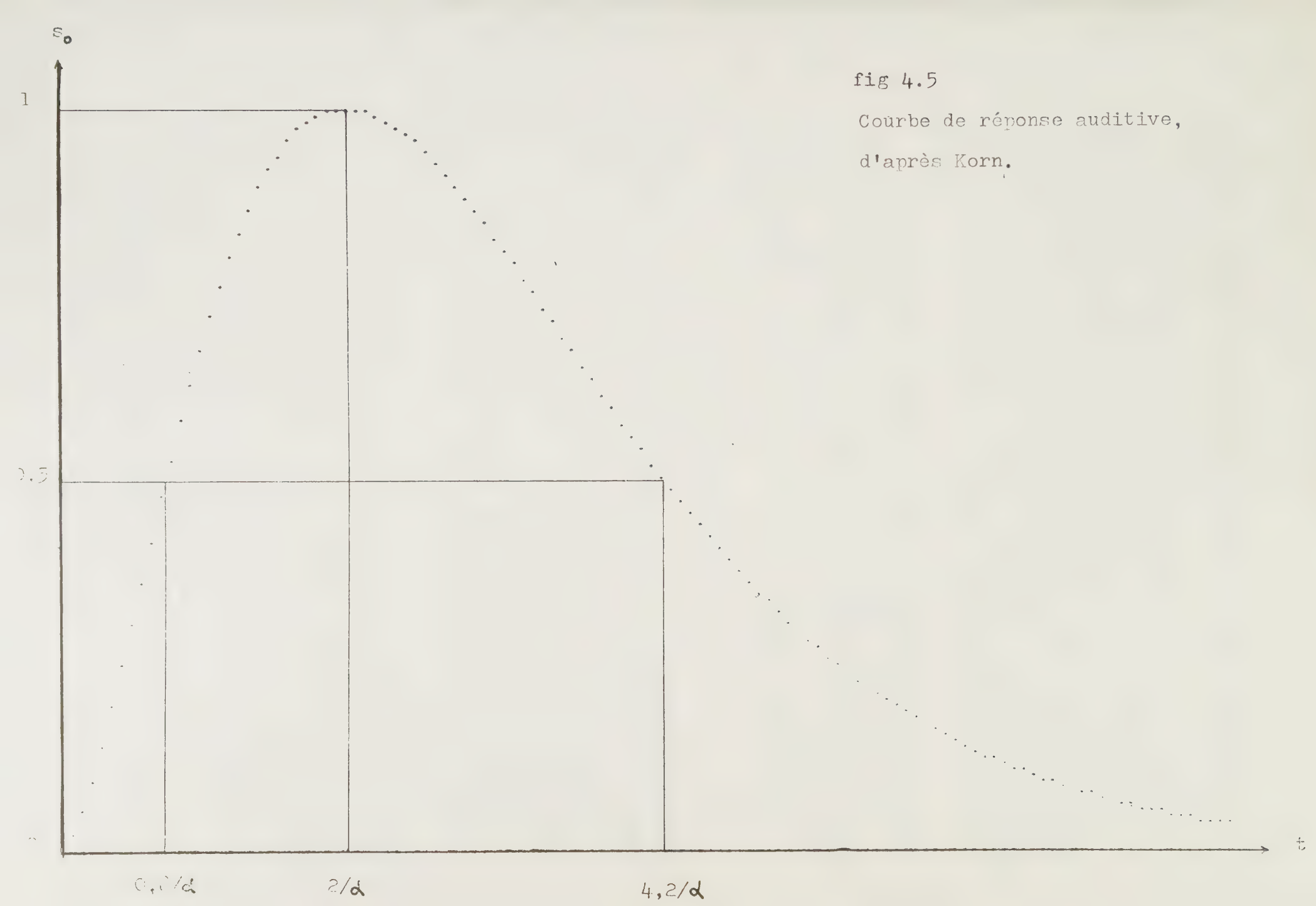


fig 4.5

Courbe de réponse auditive,

d'après Korn.

affectée du facteur de normalisation $(\alpha e/2)^2$ qui donne à son maximum la valeur unité.

Pour déterminer la réponse auditive dans une salle, il y a lieu de combiner le signal atteignant l'oreille, c'est-à-dire la fonction $\Delta w(t)$ déterminée précédemment, et la fonction $s_0(t)$ ci-dessus. Sans doute la sensation auditive est-elle malaisée à déterminer : elle dépend de certains paramètres que nous négligeons ici. Nous imaginerons un récepteur possédant certaines caractéristiques de l'oreille et se prêtant bien au calcul, quitte à interpréter ultérieurement les résultats en tenant compte des caractères négligés.

Nous imaginerons un récepteur linéaire, c'est-à-dire fournissant des quantités d'électricité proportionnelles aux énergies reçues. Cet effet n'est pas nécessairement instantané; nous attribuerons au récepteur la courbe de réponse $s_0(t)$ à une excitation instantanée. Grâce à sa linéarité, il y a lieu d'utiliser le produit de convolution; la réponse à une excitation $\Delta w(t)$ est :

$$P(t) = (S_0 * \Delta w)(t)$$

On schématise parfois la fonction $s_0(t)$ par un rectangle, dont la largeur s'appelle alors "durée d'intégration de l'oreille". La largeur du rectangle encadrant à mi-hauteur la courbe de la figure 4.5 est $3,4/\alpha$, ce qui, pour la valeur courante $\alpha = 100$, donne une largeur de 34ms. Le traitement est alors très simplifié; il rejoint celui d'une impulsion rectangulaire, dans lequel il suffit de donner à τ la valeur ci-dessus. La réponse ainsi calculée présente des discontinuités dues aux bords verticaux du rectangle, tandis que celle obtenue avec la fonction $s_0(t)$ précédente est beaucoup plus lisse.

Dans le cas général où le signal émis par S n'est plus instantané, mais se distribue suivant une loi $s(t)$, nous avons dit plus haut que la réponse de la salle est $(s * \Delta w)(t)$. C'est cette fonction qui sert d'excitation à notre récepteur électroacoustique hypothétique; la réponse du récepteur est alors $(s_0 * s * \Delta w)(t)$. Grâce à la commutativité du produit de convolution, la fonction exprimée ci-dessus peut être calculée de plusieurs façons différentes; on choisira la plus commode.

e) Remarque sur la réponse en fréquence.-

Indépendamment de notre préoccupation principale, qui est l'utilisa-

tion des échogrammes pour déterminer l'intelligibilité de la parole, on pourrait penser à utiliser la propriété bien connue, que la réponse en fréquence (transmittance) est la transformée de Fourier de la réponse impulsionnelle.

En effet, soit $s(t) = e^{i\omega t}$ le signal émis par la source S. La relation (12) donne :

$$P(t) = \mathcal{Z}(\omega) e^{i\omega t} \quad \text{avec} \quad \mathcal{Z}(\omega) = \frac{1}{2\pi} \sum_k w_k e^{-i\omega t_k}$$

Dans l'expression de $z(\omega)$ on peut rétablir les longueurs $R_k = ct_k$ des rayons et utiliser la longueur d'onde $\lambda = 2\pi c/\omega$:

$$\mathcal{Z} = \frac{1}{2\pi} \sum_k w_k \, e^{-2i\pi R_k/\lambda}$$

Quelle peut être la signification d'une réponse en fréquence calculée par un procédé qui néglige par principe les phénomènes physiques liés à la nature ondulatoire de l'énergie sonore?

Aucune réponse ne peut être donnée avant de confronter calcul et expérience.

4.3.6.- Application.-

Nous avons utilisé les méthodes exposées ci-dessus dans le cas de deux salles qui existent au LMA, tout au moins à l'état de maquette, et où il était possible de faire des vérifications expérimentales.

La première salle envisagée est un parallélipipède rectangle de 22,9m de longueur, 13,6m de largeur et 11,9m de hauteur. Le libre parcours moyen d'un rayon sonore y est de 20m. Cette salle existe à l'état de maquette à l'échelle 1/4.

La seconde salle est un heptaèdre sans faces parallèles qui sert de salle réverbérante; ses murs sont verticaux, le plafond est oblique; le libre parcours moyen est de 4m (voir Fig.4.13).

Pour la salle rectangulaire, nous donnerons deux exemples de position

du récepteur pour une situation de l'émetteur :

- Source au milieu d'un petit côté, récepteur au milieu du petit côté opposé.
- Source au milieu d'un petit côté, récepteur près d'un angle opposé.

Ces deux exemples sont traités par la méthode des images, particulièrement commode dans le cas d'une géométrie aussi simple. Les courbes obtenues sont "exactes", les coordonnées des sources images sont obtenues directement par une seule formule.

Dans le cas de la salle heptaédrique, pour une seule position de l'émetteur et du récepteur, nous avons confronté les deux méthodes de calcul. Le sol a été supposé parfaitement absorbant, la source isotrope, le coefficient de réflexion des parois indépendant de l'incidence et égal à 0,9. Le récepteur plan est un carré horizontal de 1m de côté et le récepteur sphérique une sphère de rayon r=0,25m située à 1,20m au-dessus du sol, la hauteur de la source étant de 2m.
Nous donnons comme exemples préférentiels les courbes P(t) obtenues pour τ = 34ms, qui correspondent, comme nous l'avons signalé, à une intégration auditive au découpage abrupt.

a) Salle rectangulaire.-

Dans ce cas simple, la disposition des sources images se réduit à deux couches, la première étant située au niveau de la source originale, la seconde étant au niveau de son image dans le plafond.

Dans le premier cas, on a affaire à une disposition de la figure présentant un plan de symétrie passant par l'émetteur et le milieu du pavé récepteur. La Figure 4.6, relative à un récepteur plan, présente la courbe P(t) obtenue pour τ = 34ms; son allure, assez découpée, présente de nombreux paliers locaux.

Au contraire, dans le cas suivant, la figure ne présente plus de symétries, les contributions des sources images seront beaucoup plus échelonnées dans le temps, les cas de coïncidence n'étant que fortuits. La courbe de la figure 4.7, également obtenue pour τ = 34ms, présente un aspect plus arrondi.

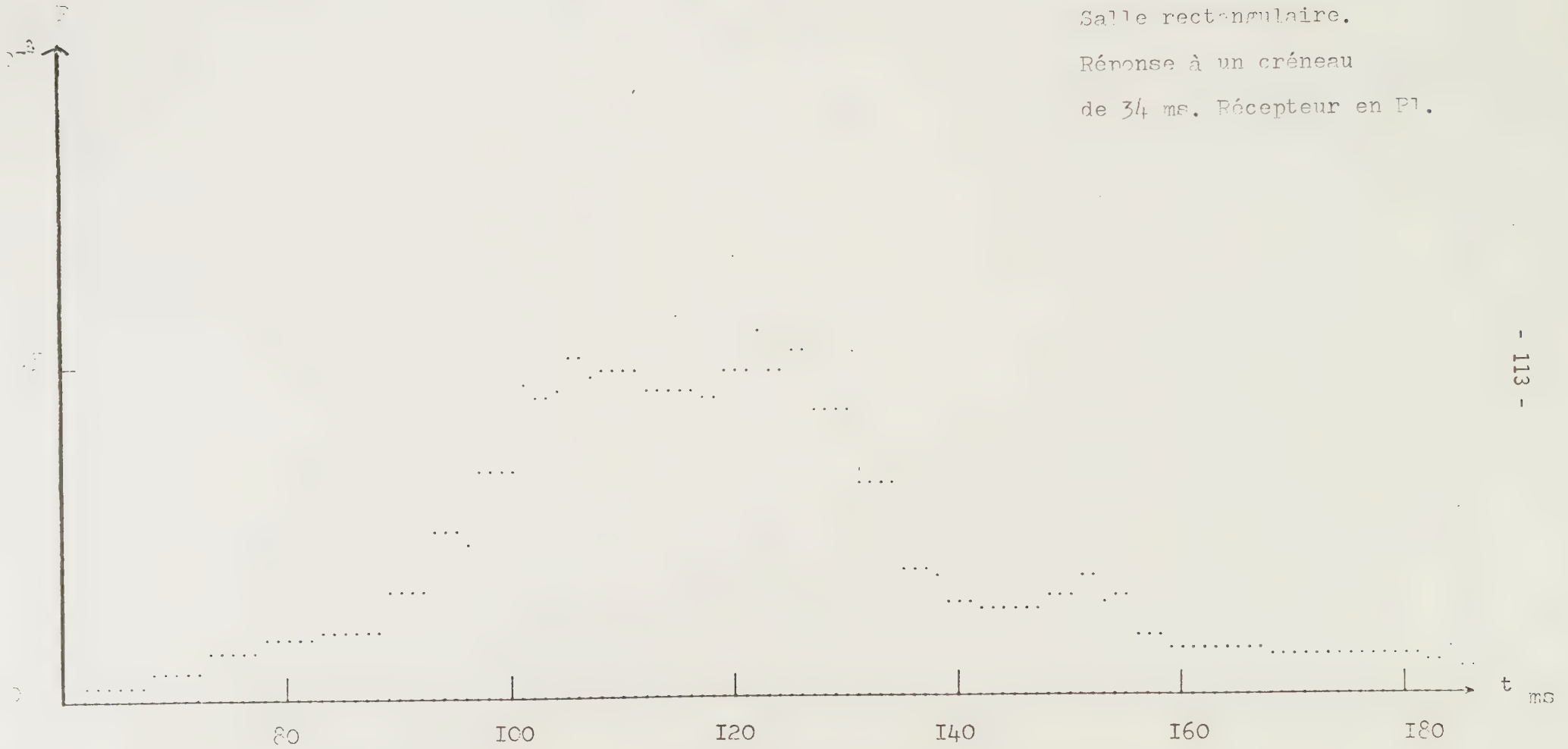
En utilisant une impulsion de 1ms de durée, on peut obtenir avec une bonne approximation l'échogramme de la salle considérée, dans les deux situations précédentes. On remarque dans le premier cas (Fig. 4.8) un regrou-

fig 4.6

Salle rectangulaire.

Réponse à un créneau

de 34 ms. Récepteur en P1.



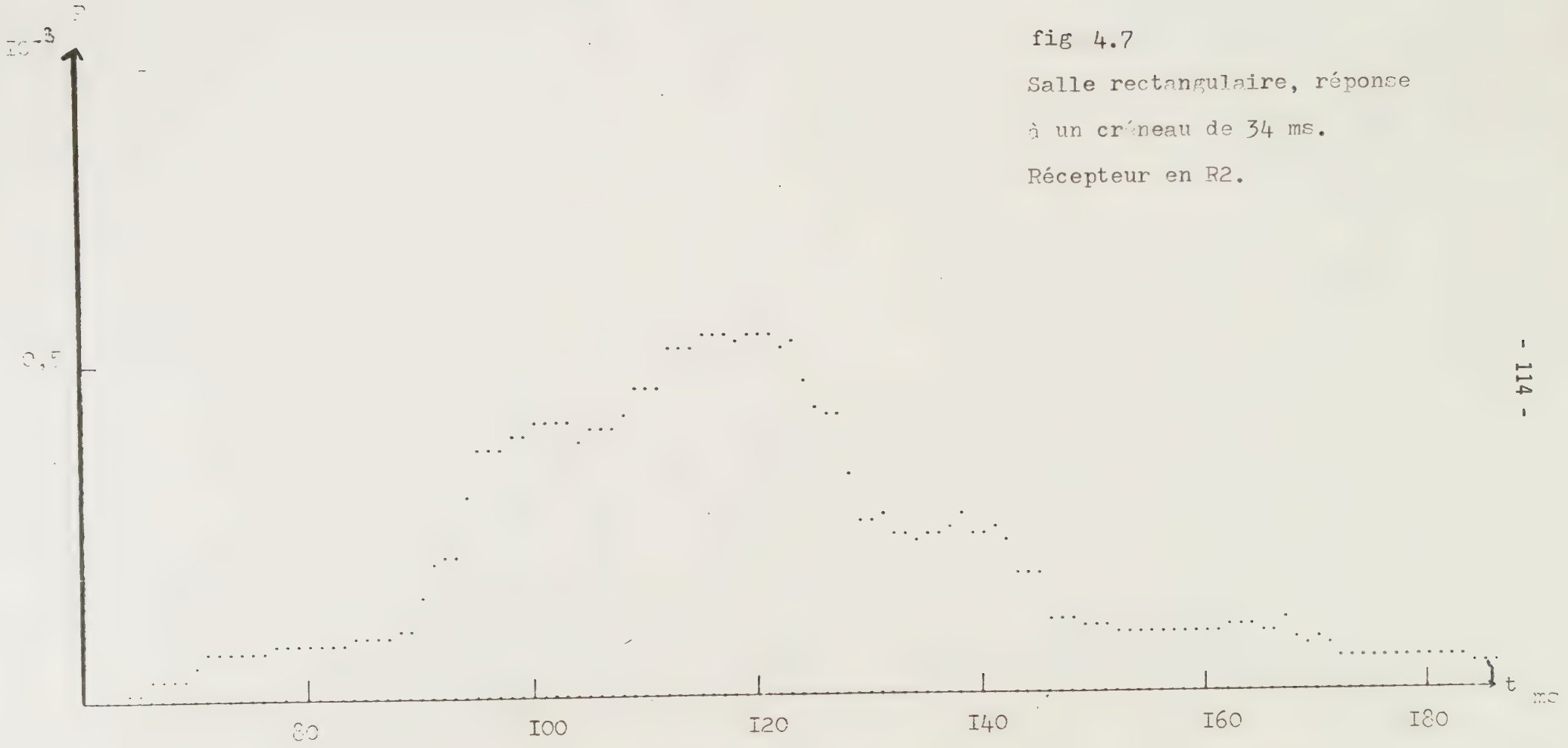


fig 4.7

Salle rectangulaire, réponse à un créneau de 34 ms.

Récepteur en R2.

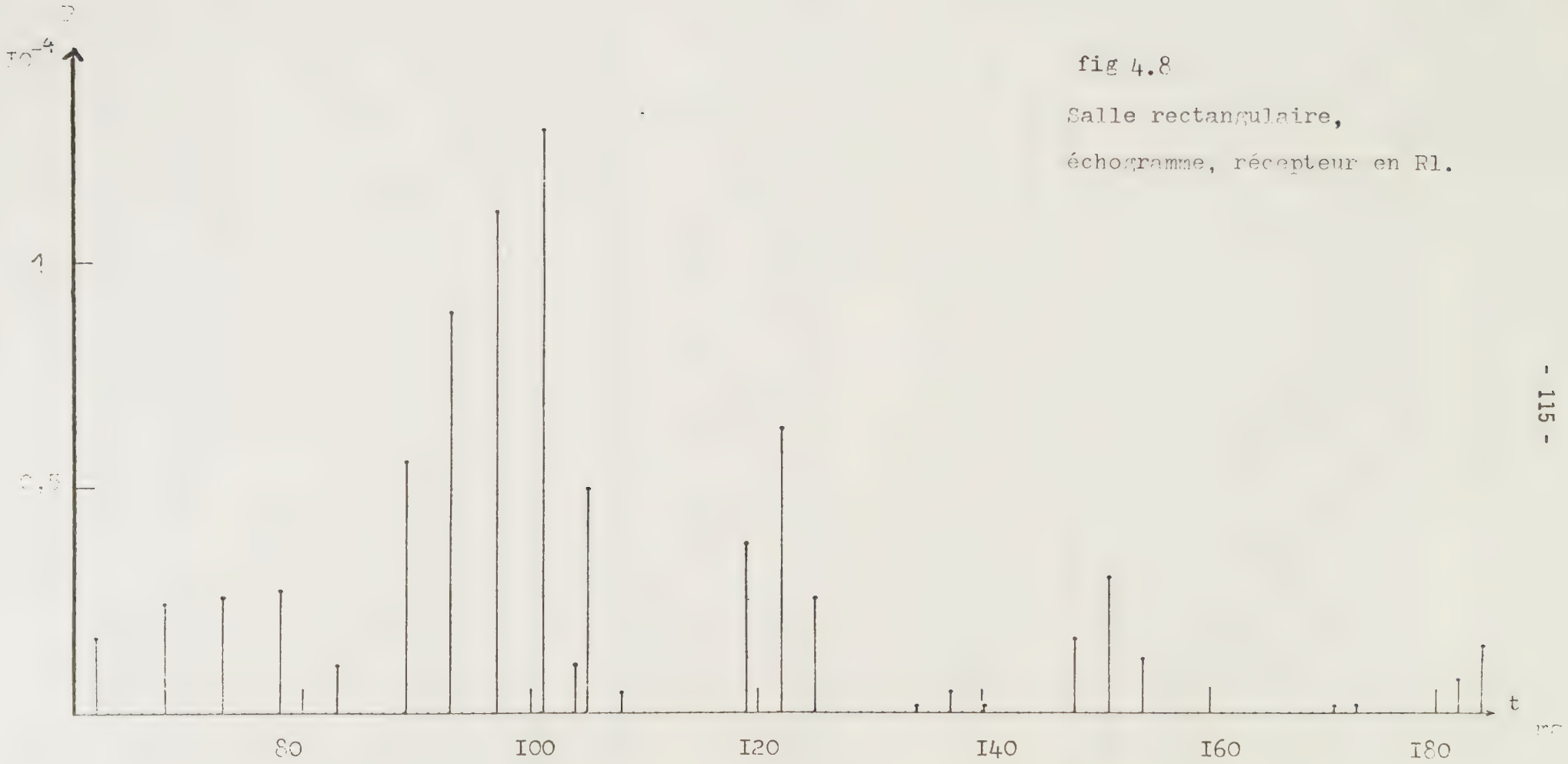


fig 4.8

Salle rectangulaire,

échogramme, récepteur en R1.

pement régulier des échos, dans le second cas (Fig. 4.9) une répartition beaucoup plus étalée.

Dans des conditions voisines de celles de la figure 4.6, mais en supposant le récepteur sphérique, nous avons calculé la courbe P(t) de la Fig. 4.10. Les discontinuités se produisent grosso modo aux mêmes moments que dans le cas de la Fig. 4.6, mais l'intensité des sources éloignées est nettement augmentée par rapport au cas du récepteur plan, étant donné que le récepteur sphérique offre à toutes les sources images une même surface apparente. Schématiquement, on peut dire que dans le cas du récepteur plan la décroissance de l'énergie s'effectue suivant une loi du type $e^{-\beta R}/R^2$, étant donné que la couronne où se trouvent localisées les sources images intervenant à un moment donné a une surface proportionnelle à R, la contribution des sources images étant en $e^{-\beta R}/R^3$.
Au contraire, dans le cas du récepteur sphérique, la contribution des sources images est en $e^{-\beta R}/R^2$, aussi la loi de décroissance de l'énergie sonore se fera-t-elle en $e^{-\beta R}/R$ (β est un coefficient positif).

Courbes de réponse pseudo-auditive.

Comme nous l'avons vu plus haut, on peut raffiner le calcul en utilisant comme courbe d'intégration de l'oreille, non pas une fenêtre rectangulaire au découpage abrupt, mais une courbe au profil adouci :

$$s_o(t) = t^2 e^{-\alpha t} \quad \text{avec} \quad \alpha = 100 s^{-1}$$

Les deux courbes P(t) ont été calculées dans les situations du récepteur et de la source précédemment définies (Fig. 4.11 et 4.12).

Comme on pouvait s'y attendre, les courbes obtenues sont bien lissées par rapport à celles relatives à des enveloppes rectangulaires : les discontinuités que l'on avait relevées précédemment ne sont pas perçues par l'oreille. Cependant, les profils de croissance et de décroissance de l'énergie reçue montrent, d'une courbe à l'autre, des différences notables, mettant en évidence le fait que l'oreille entendra différemment, d'un endroit à l'autre de la même salle.

b) Salle heptaédrique.- Méthode des images.-

La salle considérée étant de petites dimensions, on est amené rapidement à calculer les sources images jusqu'à un ordre assez élevé. La Fig. 4.14 montre les courbes cumulatives calculées en allant respectivement aux

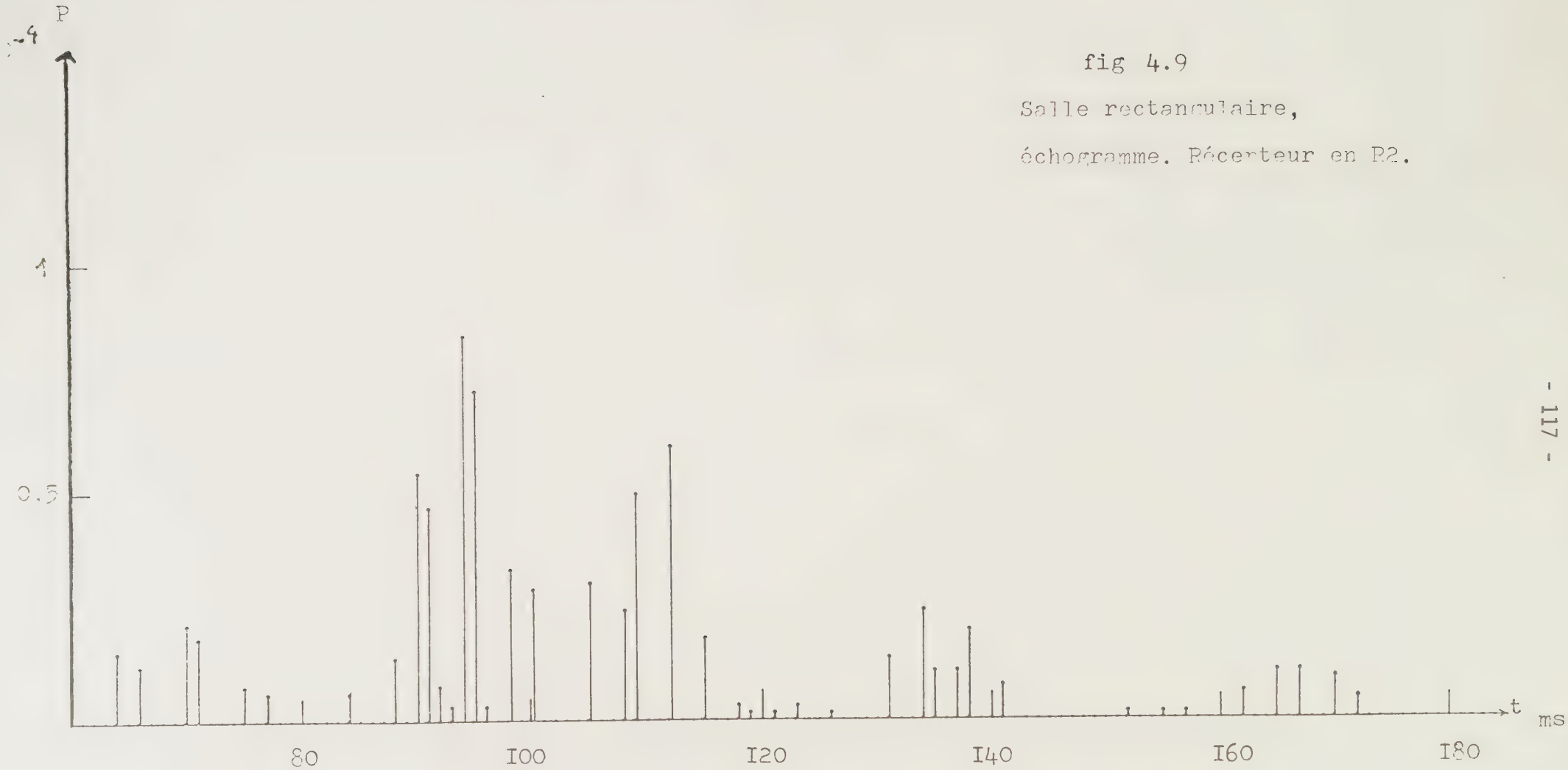


fig 4.9

Salle rectangulaire,

échogramme. Récepteur en R2.

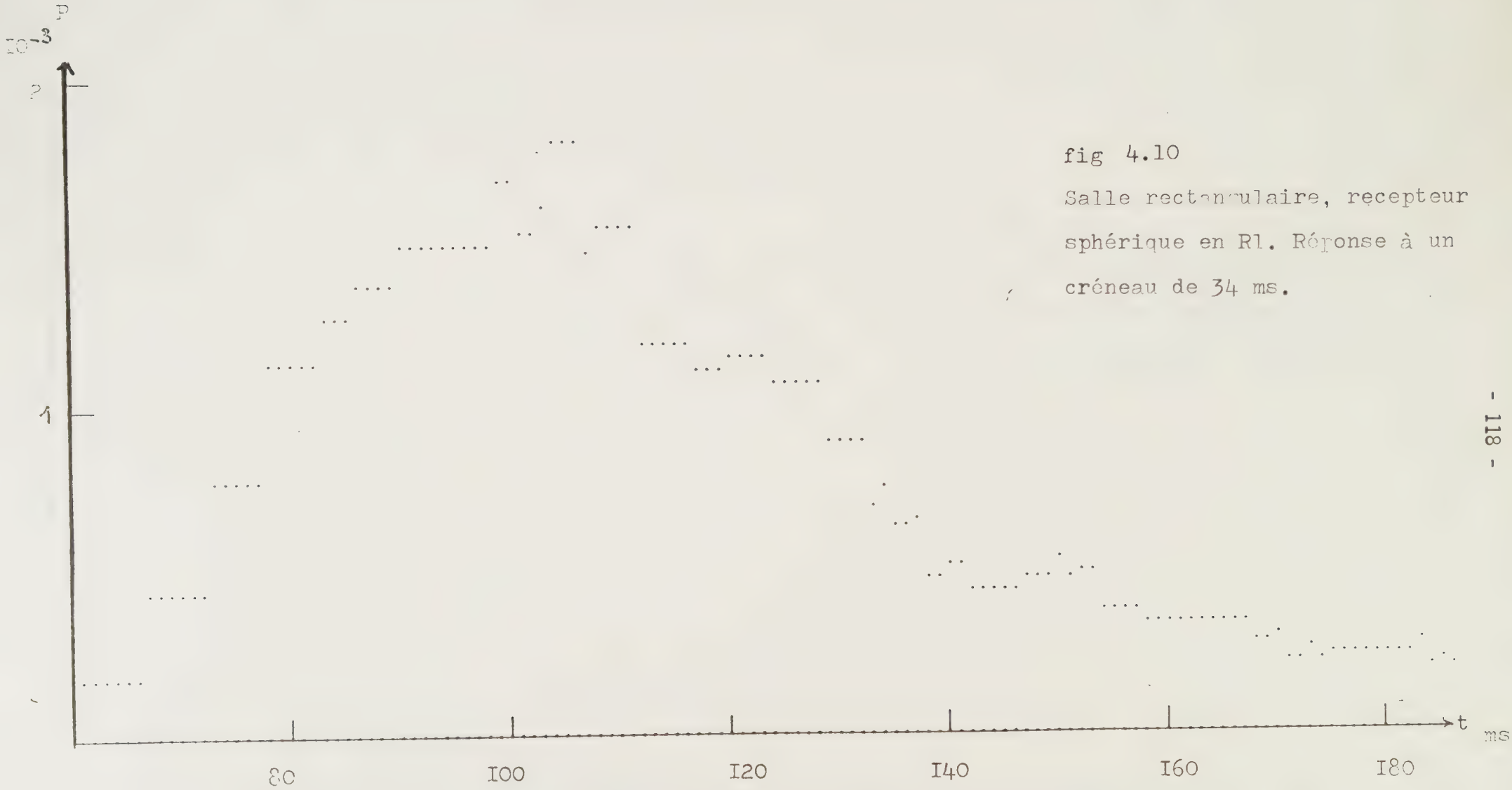


fig 4.10
Salle rectangulaire, recepteur sphérique en R1. Réponse à un créneau de 34 ms.

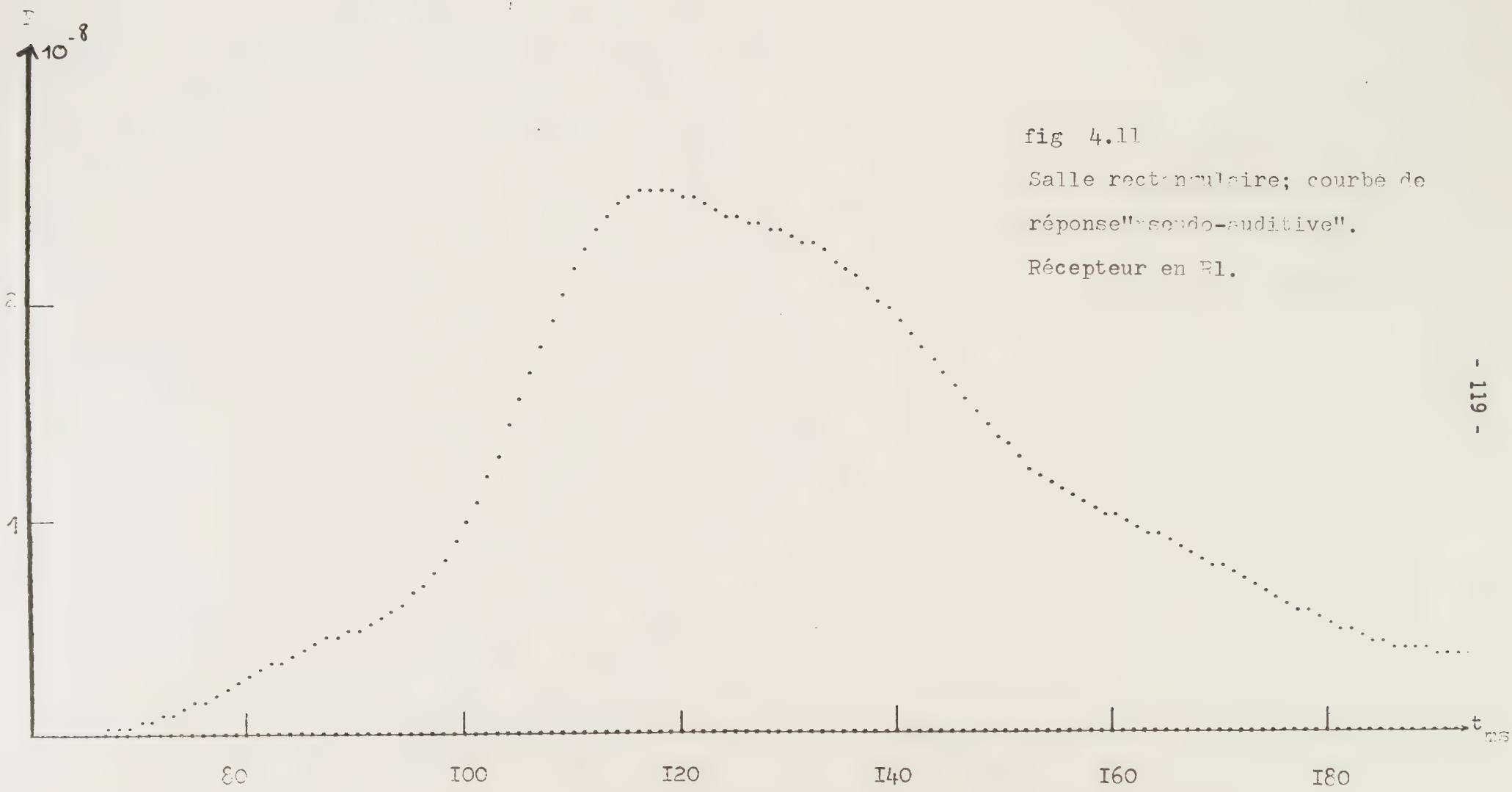


fig 4.11

Salle rectangulaire; courbe de réponse "pseudo-auditive".

Récepteur en R1.

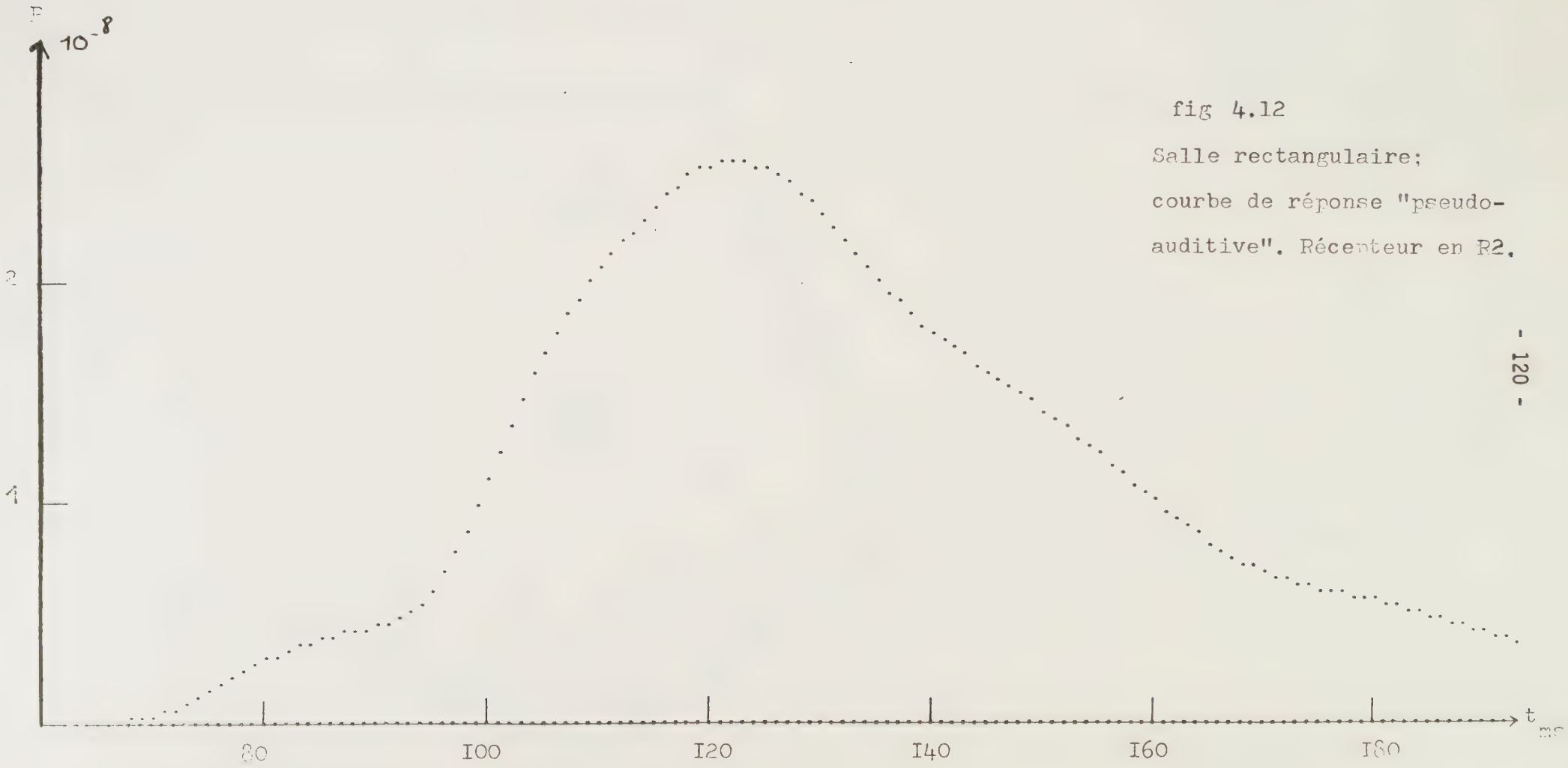


fig 4.12

Salle rectangulaire;
courbe de réponse "pseudo-auditive". Récepteur en R2.

Salle heptaédrique

coupe

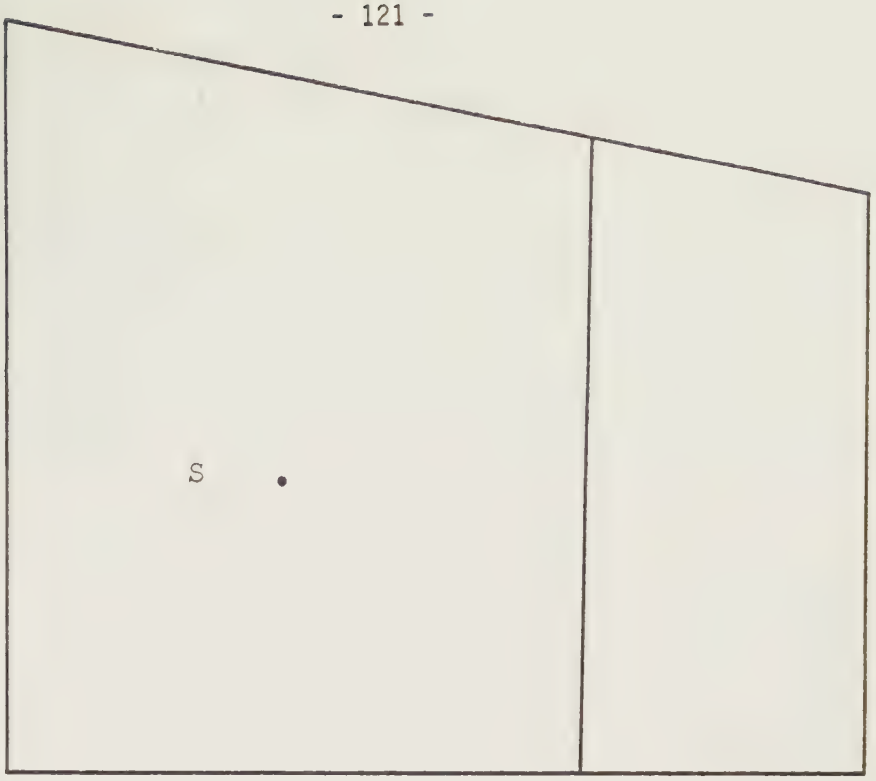


Fig 4.13

plan

1 m

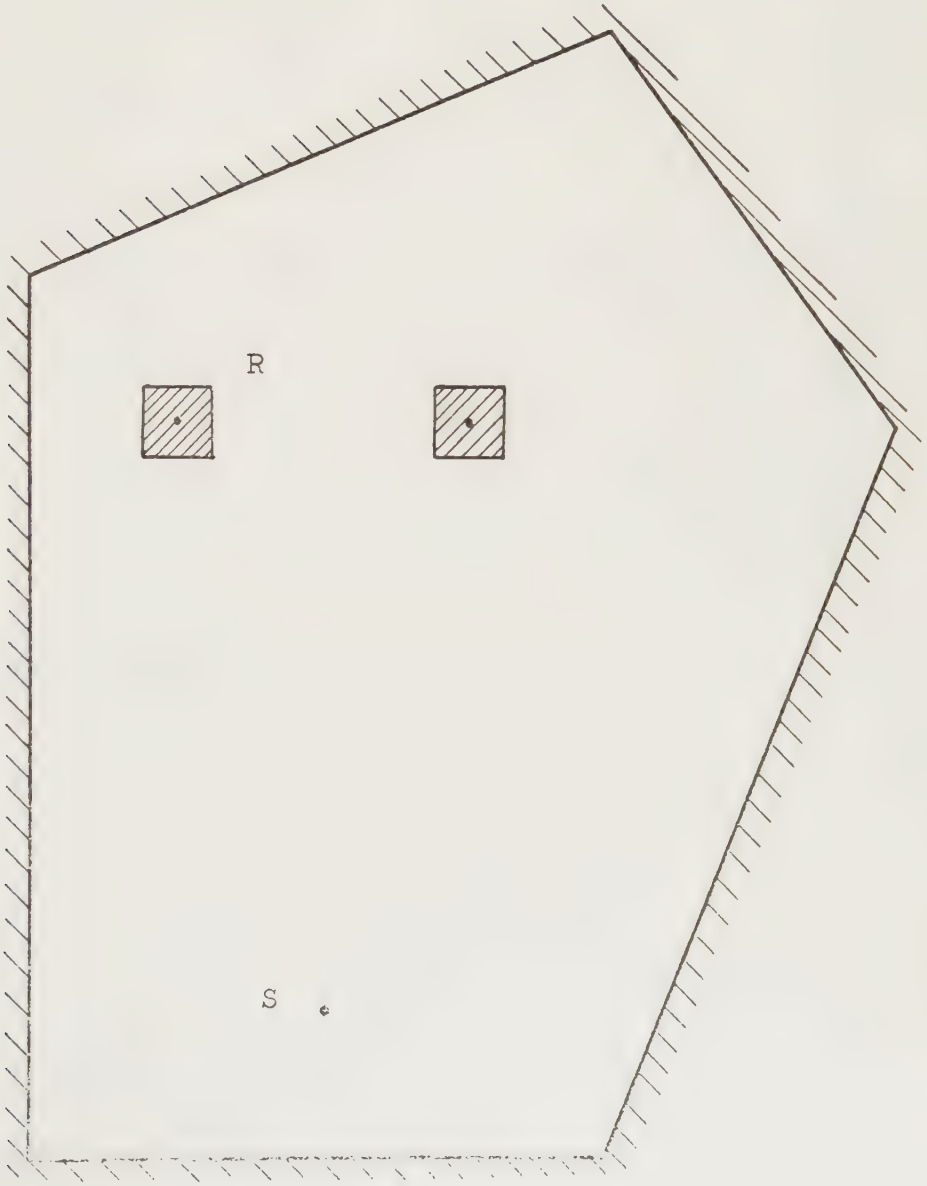
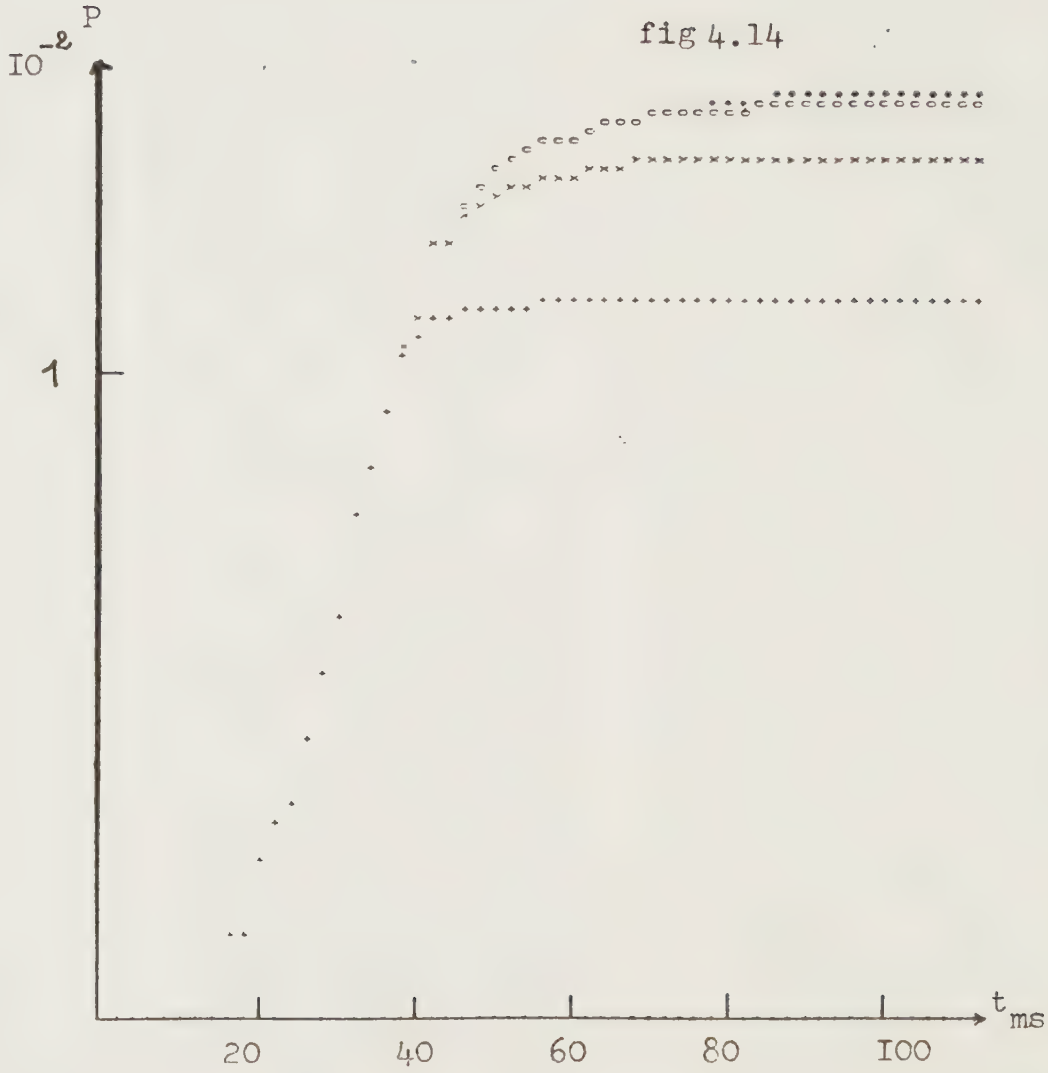


fig 4.14



Salle heptaédrique,
méthode des images.
Courbes cumulées limitées
aux 4°, 5°, 6° et 7° ordres.

4°, 5°, 6° et 7° ordres.

En comparant les courbes limitées aux 4° et 5° ordres, on voit que la courbe limitée au 4° ordre est exacte pour $t \leqslant 38ms$, de même, l'examen comparé des courbes limitées aux 5° et 6° ordres montre que la courbe limitée au 5° ordre est exacte pour $t \leqslant 46ms$.
La Fig. 4.15 montre la courbe obtenue pour $\tau = 34ms$, la Fig. 4.16 le résultat adouci par la fonction $s_o(t)$.

- Remarque sur le temps d'arrivée des sources images.-

La courbe d'énergie cumulée est obtenue en sommant les contributions des différentes sources images qui arrivent après la source primaire. Ces contributions sont échelonnées dans le temps. On peut envisager le nombre de sources images N dont les contributions sont arrivées en deçà du temps t , pour diverses valeurs de ce temps. Dans le cas de la salle heptaédrique et pour la situation envisagée de la source et du récepteur, cette distribution est donnée par la courbe N(t) de la figure 4.17.
Il peut paraître intéressant de comparer cette courbe à celle :

$$N_I = 2\pi (ct)^2 / S$$

qui correspondrait au nombre de sources images dont la contribution atteint le récepteur en deçà du temps t, dans le cas d'une salle rectangulaire de même volume V que la salle heptaédrique. Cette salle aurait un sol parfaitement absorbant de surface S égale à celle du sol de la salle heptaédrique, si bien que les sources images sont réparties sur deux couches de surface $\pi (ct)^2$, l'une étant située au niveau de la source réelle, l'autre au niveau de son image dans le plafond. Bien entendu, ce dénombrement est d'autant plus exact que le rayon ct est plus grand. A titre indicatif, on a également porté la répartition :

$$N_o(t) = 2\pi (ct)^3 / 3V$$

ce qui supposerait les images réparties régulièrement à l'intérieur d'une demi-sphère de rayon ct, ce qui est manifestement faux dans le cas d'une salle heptaédrique, où les images sont d'autant plus rares qu'elles sont plus élevées au-dessus du niveau du sol.
Les variations de N(t) et de $N_1(t)$ sont assez voisines pour des délais assez faibles (inférieurs à 55ms). Par contre, au-delà de ce délai, le nombre N(t) de sources images qui interviennent semble beaucoup plus faible pour la salle heptaédrique que pour la salle rectangulaire.

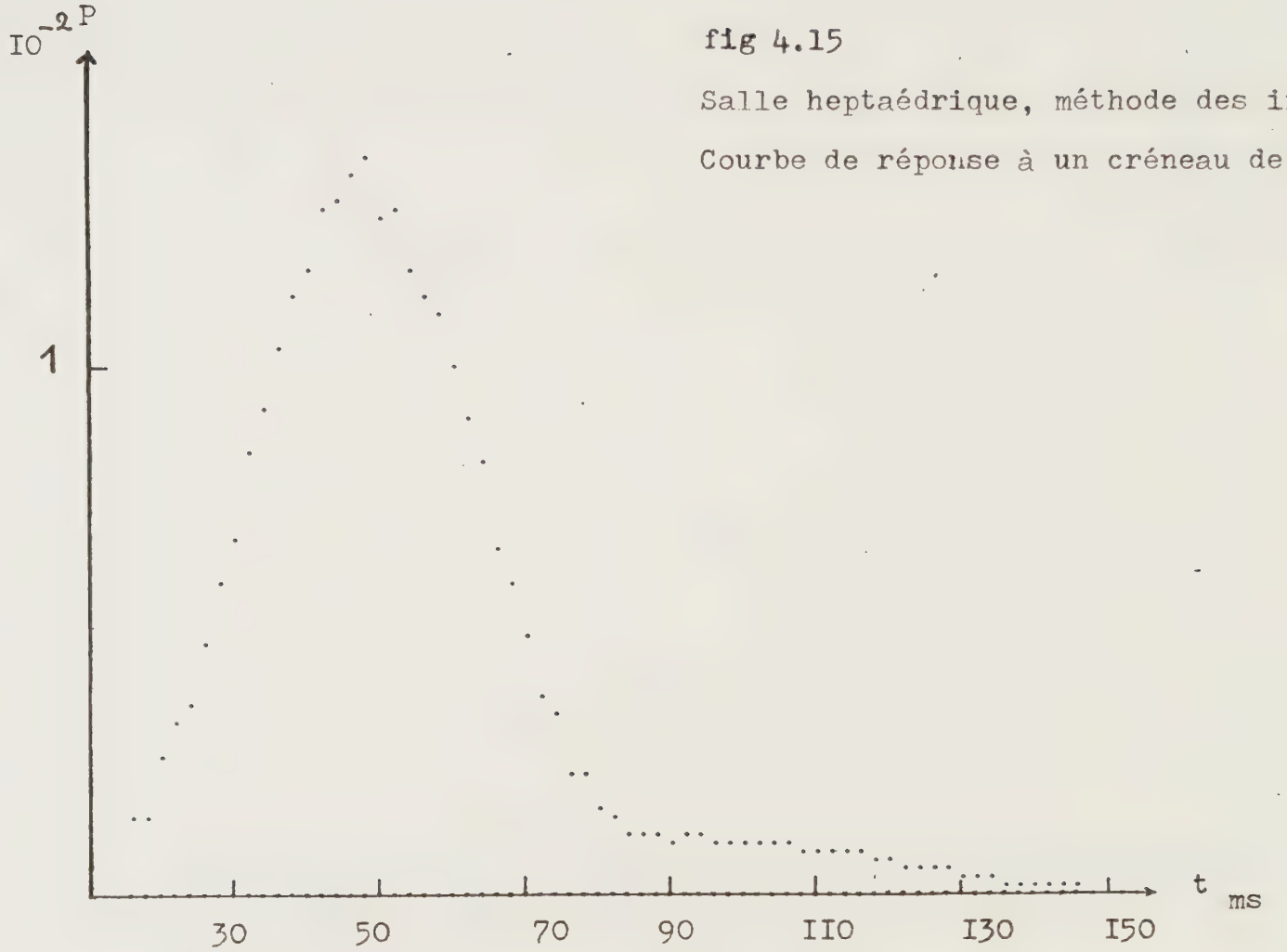


fig 4.15

Salle heptaédrique, méthode des images.

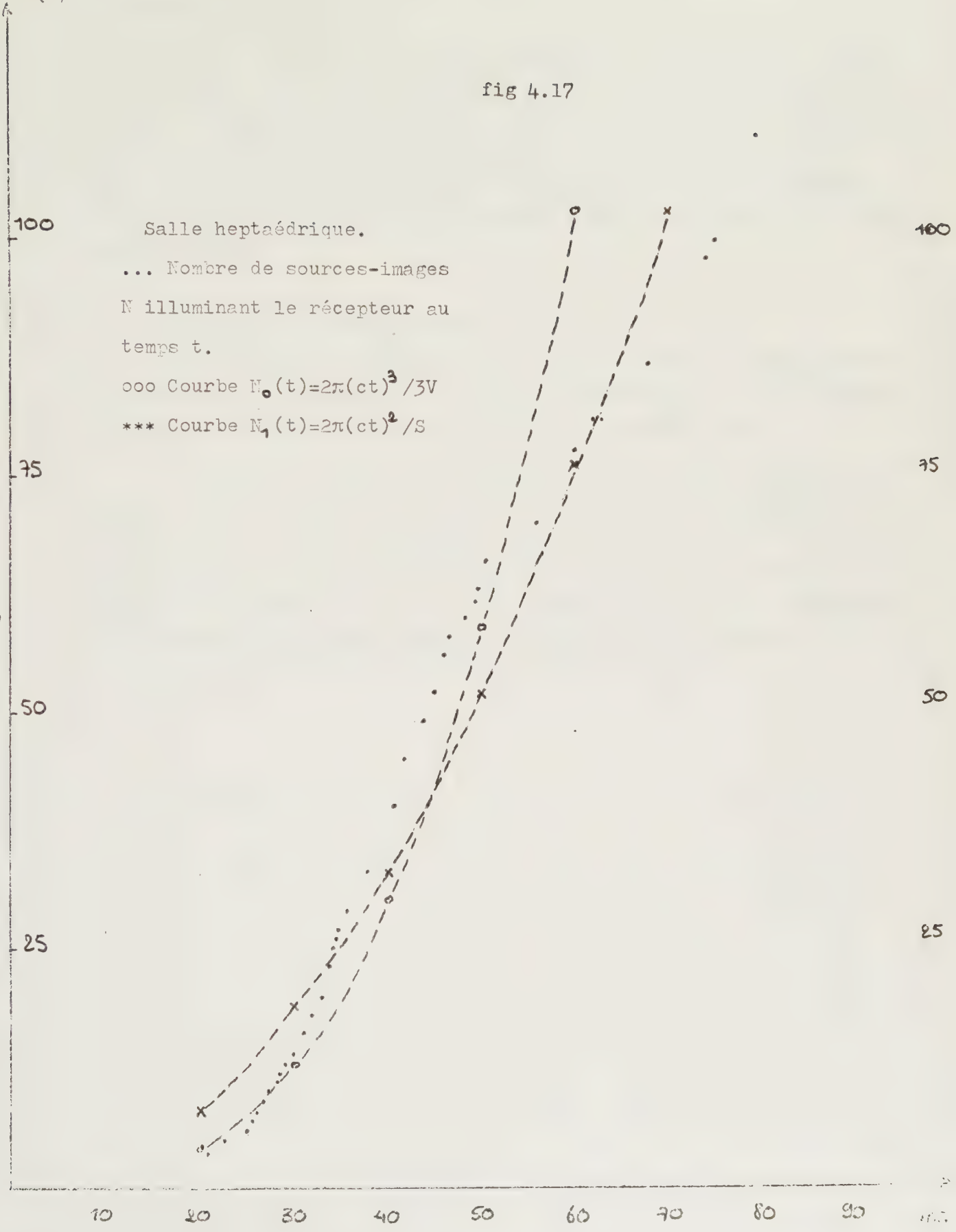
Courbe de réponse à un créneau de 34 ms.

fig 4.16

Salle hémisphérique, méthode des images.

Courbe de réponse "pseudo-auditive".

fig 4.17


N(t)
100
75
50
25
10
20
30
40
50
60
70
80
90
Salle heptaédrique.
... Nombre de sources-images N illuminant le récepteur au temps t.
ooo Courbe $N_0(t)=2\pi(ct)^3/3V$
*** Courbe $N_1(t)=2\pi(ct)^2/S$

Méthode des rayons.

La Figure 4.18 montre les courbes cumulatives W(t) obtenues, en émettant 1000 rayons, pour deux tirages différents. La Figure 4.19 montre les courbes analogues aux précédentes, et relatives à une émission de 4000 rayons. La courbe relative à τ = 34ms (Fig. 4.20) et la courbe pseudo auditive (Fig. 4.21), obtenues également pour une émission de 4000 rayons montrent une très bonne concordance avec les courbes similaires obtenues par la méthode des images.

c) Efficacité des deux méthodes.-

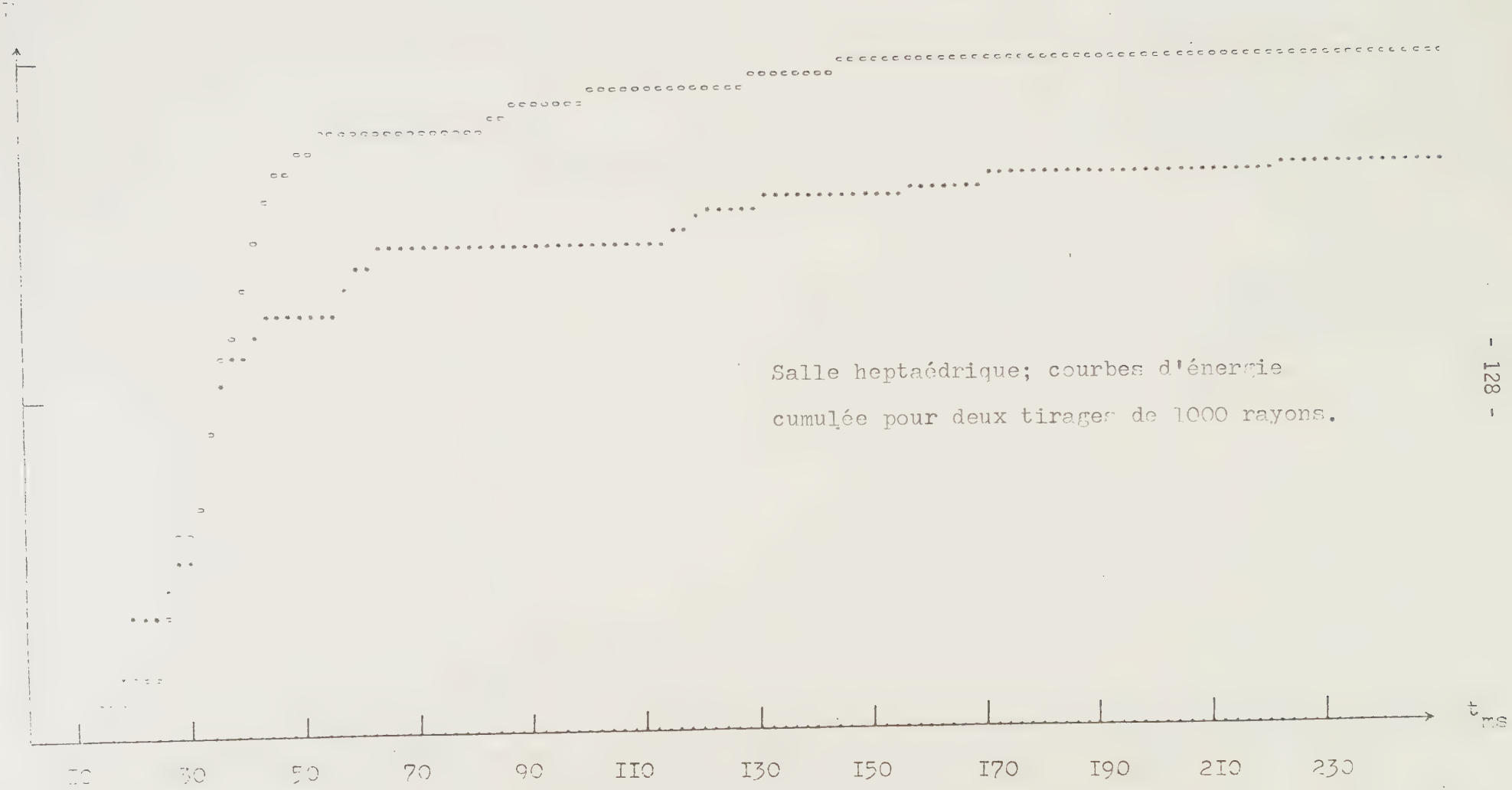
Nous regroupons ci-dessous les principaux résultats acquis lors de l'application des deux méthodes à la salle heptaédrique, en ce qui concerne les erreurs de méthode et les temps de calcul sur UNIVAC 1106.

Le tableau I indique, pour chaque valeur de l'ordre maximal m des sources images, les temps t_1 et t_2 correspondant respectivement à la première et à la dernière des sources images d'ordre m, le pourcentage % qui représente l'énergie des sources images d'ordre m par rapport à l'énergie cumulée jusqu'à l'ordre m , et le temps de calcul T_c, comprenant le recensement des sources images, la construction de la fonction w(t) et son utilisation au tracé de diverses courbes P(t).

m	t_1(ms)	t_2(ms)	%	T_c
4	38	78	I8,I	I6s
5	46	92	6,35	I9s
6	74	II2	I,7	20s
7 et 8	82	I44	3,3	47s
9 et IO	94	I78	2,I5	4mn48s

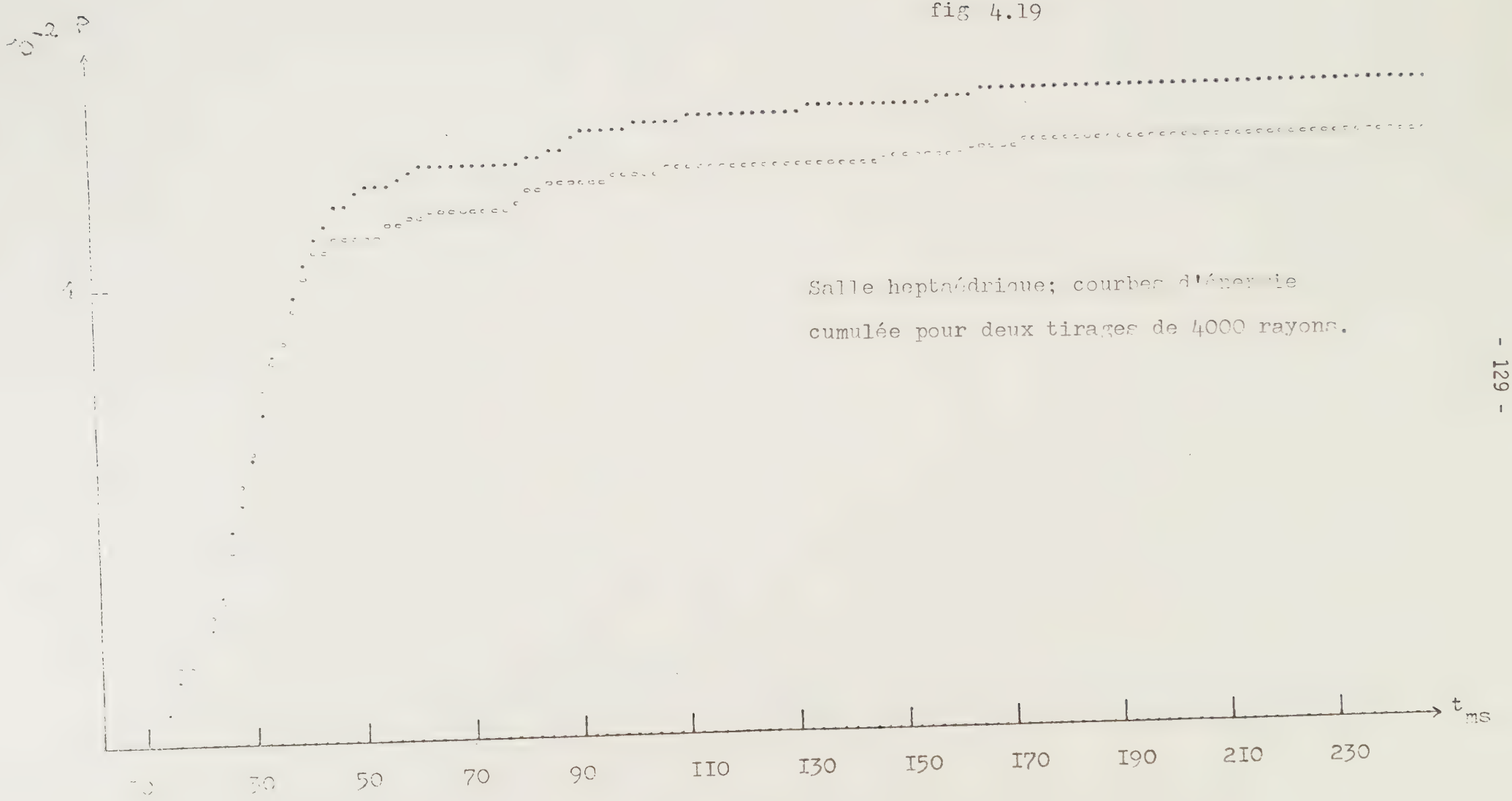
Tableau I

fig 4.18



Salle heptaédrique; courbes d'énergie cumulée pour deux tirages de 1000 rayons.

fig 4.19



Salle heptaédrique; courbes d'énergie cumulée pour deux tirages de 4000 rayons.

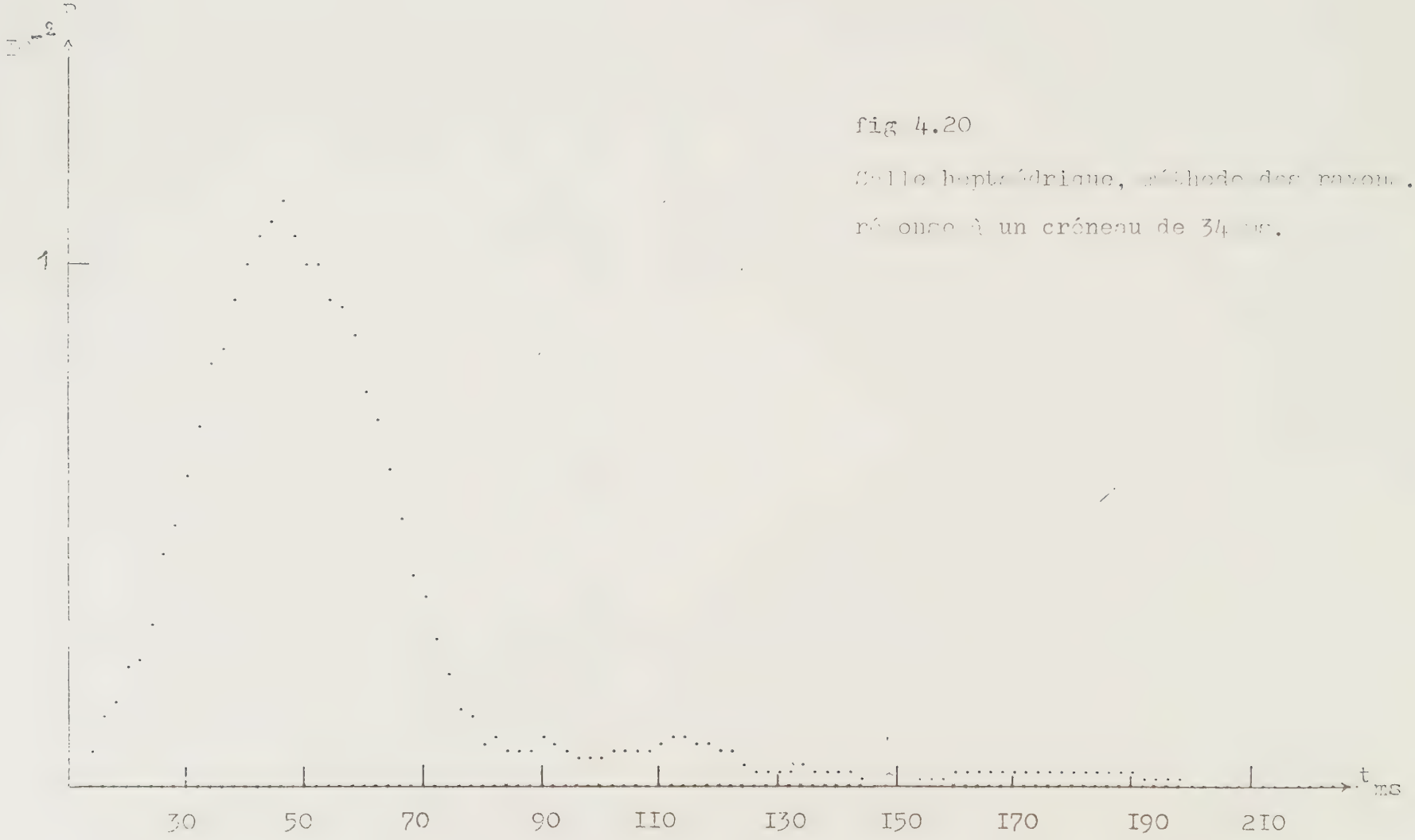


fig 4.20

Cellule hepta édrique, méthode des rayons.

réponse à un créneau de 34 ms.

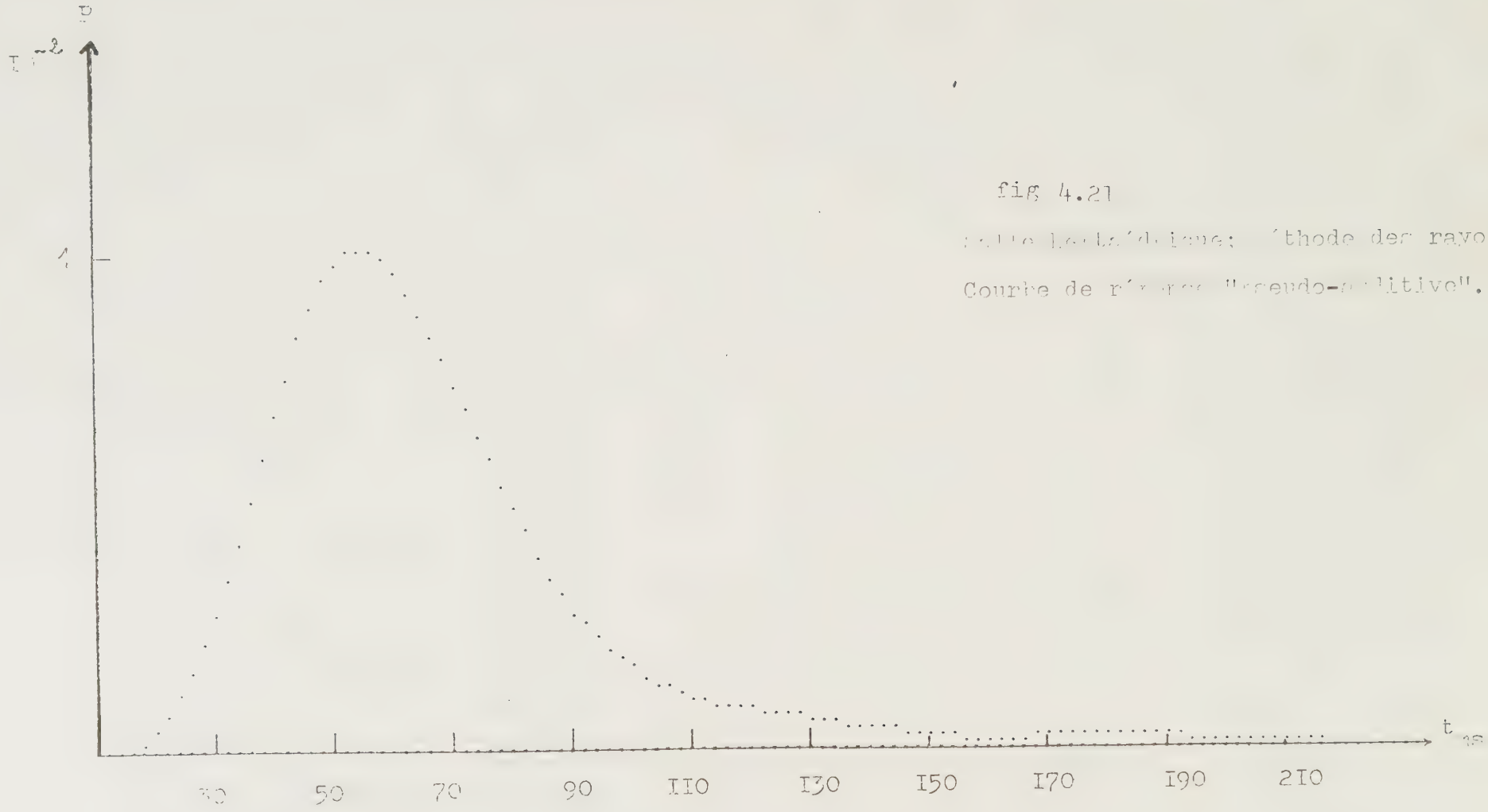


fig 4.21

Courbe de réponse "pseudo-cumulative".

Le tableau II indique, pour chaque valeur du nombre N de rayons discrets émis, le nombre n de rayons parvenue au récepteur, l'écart type σ estimé sur l'énergie reçue, et le temps de calcul Tc , comprenant l'essai des N rayons, suivis des mêmes applications qu'avec l'autre méthode. Les résultats pour n et σ sont les moyennes, pour chaque valeur de N, de six essais distincts. Le temps de calcul T_c est relatif à l'ensemble de quatre tirages par valeur de N.

N	n	σ %	T_c	σ' %
1000	23	18,5	1mn39s	21,8
2000	44	7,6	2mn38s	19,2
4000	81	7,1	4mn58s	9,9
8000	180	7,6	9mn52s	5,2

Tableau II

Ces résultats étaient obtenus après un découpage préalable de la sphère en pavés émetteurs. A titre de comparaison, nous avons porté les écarts-types σ' trouvés lorsque l'on tire chacun des points D_k au hasard sur la sphère trigonométrique toute entière.

Pour un temps de calcul T_c donné, la méthode des images permet de définir deux temps t_1 et t_2 : l'erreur étant nulle en deçà de t_1 , inacceptable au-delà de t_2 , il existe une valeur t_o de t , comprise entre t_1 et t_2 pour laquelle l'erreur est égale à l'écart σ de la méthode des rayons. On pourra alors définir deux domaines de variations de t : pour $t < t_o$, la méthode des images est plus avantageuse; au contraire, pour $t > t_o$, on aura tout intérêt à utiliser la méthode des rayons. Dans le cas de l'exemple précédent, et pour un temps de calcul de 5 min. environ, la méthode des rayons s'impose pour $t > 120$ ms.

4.3.7.- Vérification expérimentale.-

Aussi curieux que cela puisse paraître, il n'existe pas dans la littérature de vérification expérimentale des échogrammes. Parmi les auteurs qui ont travaillé sur l'exploitation des échogrammes, beaucoup se sont fiés aux échogrammes enregistrés, que ce soient Niese, Lochner ou Januska. Par contre, d'autres ont effectué des calculs, mais sans aucune contre partie expérimentale, faisant entièrement confiance au résultat théorique, même basé sur des hypothèses criticables à certains égards. C'est ainsi qu'il nous est apparu nécessaire d'envisager l'un et l'autre des points de vue, et de voir comment ils peuvent se raccorder.

a) Remarques sur le spectre de l'impulsion.-

Pour obtenir un échogramme expérimental, quelle impulsion doit-on envoyer dans une salle? et d'autre part, dans la pratique, quelle impulsion peut-on envoyer dans une salle?

Nous avons supposé, dans nos calculs, que la source émettait une énergie unitaire à l'instant $t = 0$. Telle est l'hypothèse fictive qu'on ne saurait reproduire dans la réalité. Il faudra alors se résoudre à n'envisager qu'une énergie émise pendant un laps de temps assez bref, et même aussi bref que possible : cette brièveté est un élément particulièrement important si l'on veut obtenir un échogramme $\Delta w(t)$ en valeurs instantanées : en effet, on doit pouvoir obtenir un enregistrement où les différents échos soient suffisamment espacés pour qu'ils soient discernables, tout au moins au début de l'enregistrement. S'il s'agit de la courbe d'énergie cumulée $w(t)$, cette condition n'est pas indispensable,nous aurons l'occasion de revenir plus loin sur cet aspect de la question.

Revenons au problème posé par la brièveté de l'impulsion : on a obtenu des impulsions particulièrement brèves en utilisant une décharge électrique : c'est le procédé employé par Thiele dans ses travaux sur la netteté [3.27] : l'impulsion obtenue a une durée de 0,1 ms. Son spectre, d'après Thiele, tombe brutalement au-dessus de 500 Hz. C'est là que se pose, de toute évidence, le problème du spectre de l'impulsion utilisée, et des fréquences qu'elle contient, auxquelles correspondent divers coefficients de réflexion des parois : car, si dans le calcul théorique, basé sur l'acoustique géométrique, nous avons fait abstraction de la notion de fréquence, il ne

saurait en être de même pour ce qui est du problème expérimental : on est alors obligé de prendre en considération la notion de fréquence, et l'échogramme obtenu risque d'être différent suivant le type de l'impulsion envisagée. Lochner et Burger [3.31] ont utilisé soit une étincelle, dont ils ne précisent pas la durée, soit un train d'onde de 10 ms, dont la fréquence glissait de 1000 à 1200 Hz. Les résultats obtenus dans l'un et l'autre cas sont très voisins. Au contraire, Niese, [3.35] a utilisé plusieurs impulsions de spectre différent, dont la largeur s'étendait de 700 à 2300 Hz, et centrées autour de fréquences situées entre 428 et 1750 Hz. D'une façon générale, la netteté au sens de Thiele est d'autant plus faible que la fréquence centrale de l'impulsion considérée est plus faible, autrement dit, la netteté varie en sens inverse de la durée de réverbération. Si Thiele travaillait "dans l'abstrait", sans se soucier de tel ou tel type de signal acoustique, Niese a essayé de savoir quelle était l'impulsion la plus significative pour obtenir un échogramme destiné à une prévision d'intelligibilité. Le choix de Niese s'est effectué de la façon suivante : on envisage les variations d'intelligibilité en fonction du spectre d'un écho retardé de 100 ms, et soumis à un filtrage d'un octave de largeur. La diminution de l'intelligibilité est maximale lorsque la fréquence centrale du filtrage est de l'ordre de 900 Hz : on en déduit une courbe $C(\omega)$ (Fig. 4.22) l'influence du filtrage sur l'intelligibilité est maximal pour la fréquence de 900 Hz. On peut considérer que cette courbe $C(\omega)$ est le produit de la courbe $a(\omega)$ donnant le spectre de la parole par une courbe $b(\omega)$, courbe de pondération de l'effet écho, qui est ainsi la courbe $C(\omega)$ramenée à un spectre uniforme.
Ainsi, Niese choisit comme spectre de l'impulsion envoyée dans la salle la courbe $b(\omega)$ (Fig. 4.23) chaque fréquence devant avoir l'importance effective d'un écho sur la parole. Remarquons que le train d'onde choisi par Lochner et Burger sans raison apparente, tout au moins sans raison déclarée, est lui aussi centré sur une fréquence de 1000 Hz. Il semble bien que cette fréquence soit une "bonne fréquence" pour l'étude des phénomènes d'intelligibilité, située dans une région très caractéristique de la parole.

b) Remarques sur la forme de l'impulsion.-

Une question peut se poser : la courbe d'énergie cumulée $w(t)$ obtenue est-elle indépendante de l'impulsion émise, à supposer que cette impulsion garde le même spectre de fréquences, ou encore, si l'on fait abstraction du

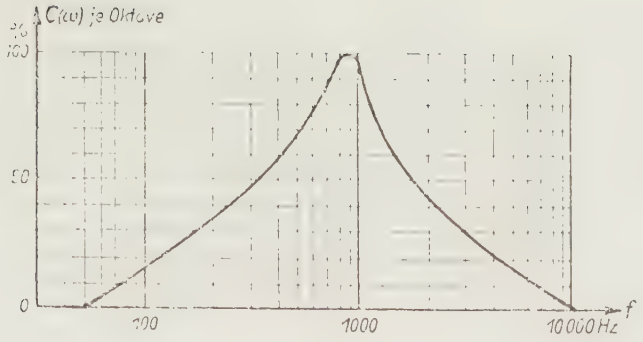


Fig. 4.22. Courbe de puissance d'un écho retardé de 10 ms en fonction d'un filtrage d'une octave en de fréquence centrale f, d'après [ESE].

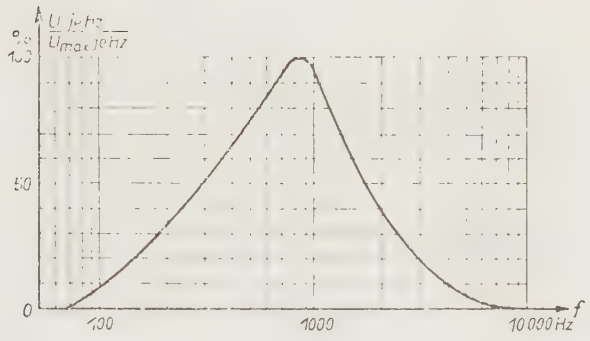


Fig. 4.23. Même courbe ramenée à un spectre uniforme.

spectre de fréquences?
Dans tous les travaux mentionnés ci-dessus, et dans le nôtre, on émet une impulsion $P_o(t)$ où P_o est une pression sonore. On recueille au récepteur une pression p(t), et pour obtenir la courbe d'énergie cumulée, on calcule l'intégrale :

$$\int_0^t p^2(t)\, dt$$

Si $p_o(t-\tau_o)$ est l'impulsion reçue au temps τ_o correspondant au son direct, et si $A_i\ p_o(t-\tau i)$ est l'image correspondant à la i-ième réflexion retardée d'un délai τ_i, et atténuée suivant le coefficient A_i , la pression totale reçue est :

$$p(t) = p_o(t - \tau_o) + \sum_i A_i\ p_o\ (t - \tau_i)$$

L'énergie cumulée au temps t est alors :

$$\int_0^t p^2(t)\, dt = \int_0^t \left\{ p_o^{\,2}\ (t - \tau_o) + {}_i\ A_i^2\ p_o^2\ (t - \tau i) \right\} dt + 2\ I(t)$$

I(t) désigne une somme d'intégrales du type $\int_0^t p(t - \tau_k)\ p(t - \tau_\ell)\ dt$ affectées de coefficients convenables.

Si l'impulsion émise est idéalement brève, c'est-à-dire si $P_o(t) = 1$ pour $t = 0$ et $P_o\ (t) = 0$ pour $t = 0$, les intégrales ci-dessus seront nulles pour $\tau_k \neq \tau_\ell$: comme l'impulsion expérimentale n'est pas idéalement brève, il risque toujours d'y avoir des chevauchements entre les diverses réflexions, et la somme I(t) ne sera pas nulle obligatoirement.

Supposons, pour fixer les idées, que le signal émis soit un train d'onde sinusoïdal de durée T. Les intégrales I(t) seront du type :

$$\int_0^{T_i} \sin \omega t.\ \sin(\omega t + \varphi_i)\, dt$$

si T_i est le temps d'interaction des réflexions d'indice k et l : bien entendu, T_i est toujours inférieur à T.

On trouve dans ce cas que I(t) est la somme de termes de la forme :

$$(T_i/2)\cos\varphi_i - (1/4\omega)\sin(2\omega T_i + \varphi_i) - (1/4\omega)\sin\varphi_i$$

avec $\omega = 2\pi f$.

Si la valeur de cette somme est faible devant celle des termes carrés, ce qui est possible, dans la mesure où φ_i peut prendre n'importe quelle valeur, on peut estimer que la courbe w(t) obtenue ne dépendra pas de la durée de l'impulsion.

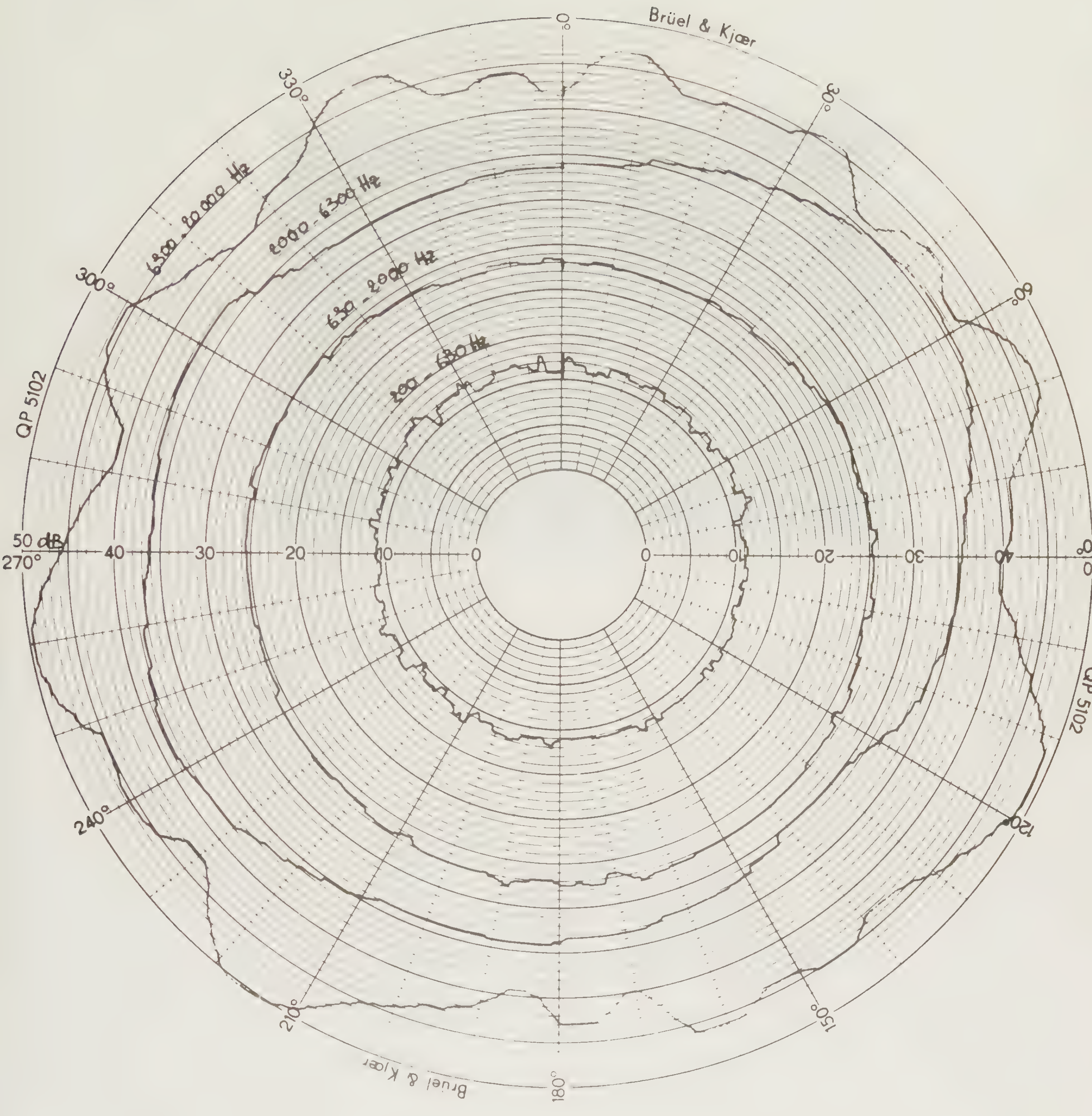
c) Montage expérimental.-

La source d'impulsions est constituée par un ensemble de six haut-parleurs montés sur une sphère en plastique. Comme nous voulions que ces haut parleurs soient très amortis, ils ont été montés avec des tampons d'ouate qui limitent le mouvement des membranes (Fig. 4.24).
Cette source a été réalisée et mise au point par Th. Angélini. Nous avons choisi un rayonnement sphérique pour la source d'émission : il eut été plus logique à certains égards, de prendre une source dont le rayonnement se rapproche le plus possible de celui de la bouche d'un locuteur réel, c'est-à-dire une tête artificielle, comme c'est le cas lors des expériences de Niese ou de Januska. Mais notre but était de vérifier expérimentalement un échogramme obtenu au moyen d'un calcul numérique, et le rayonnement de la source importait peu en l'occurence; d'autre part, il était plus facile d'obtenir un rayonnement sphérique satisfaisant, au point de vue expérimental (Fig. 4.25) plutôt que de reporter dans le programme le rayonnement obtenu avec une source du type de la tête artificielle.

Le dispositif d'alimentation est constitué par un générateur Basse Fréquence Schomand et un découpeur Général Radio, qui fournit, par l'intermédiaire d'un amplificateur à large bande, un signal rectangulaire de 0,1 à 0,2 ms de durée : la réponse de la source est donnée par la Figure 4.26 : comme on le voit, l'amortissement est très réduit. Le signal à proprement parler ne dépasse pas 2 ms. Ce qui apparaît 5 ms après le son direct est dû à un écho de la chambre sourde, provoqué par la table tournante sur laquelle était montée la source. Le spectre de cette impulsion, assimilable à une période de sinusoïde à 1000 Hz se trouve centré sur la fréquence de 1000 Hz.

Le dispositif de réception et de traitement du signal mérite une mention spéciale. Comme nous ne possédions pas de quadrateur électronique, nous avons échantillonné la courbe p(t) reçue à l'aide d'un enregistreur DIDAC 4000,

Fig. 4.24.- Source d'impulsions expérimentale -


Brüel & Kjær
0°
30°
60°
90°
120°
150°
180°
210°
240°
270°
300°
330°
QP 5102
50 dB
6300 - 20 000 Hz
2000 - 6300 Hz
630 - 2000 Hz
200 - 630 Hz
0
10
20
30
40
50

[illegible] de la source

[illegible]

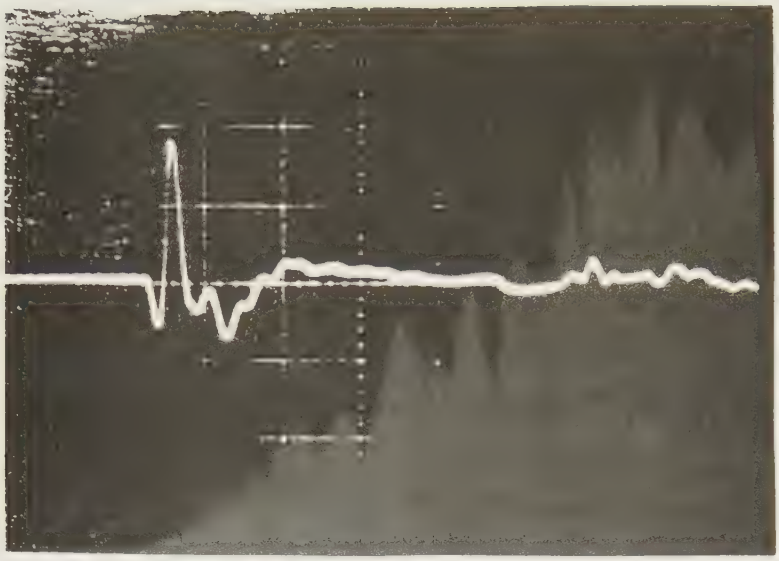


Fig. 16

Réponse de la source expérimentale à un créneau de 0,1 ms.
échelle horizontale: 10 ms par carreau.

à une fréquence assez élevée, soit 4000 Hz : cette fréquence est suffisante pour échantillonner un signal à 2000 Hz. Nous avons donc obtenu les valeurs de la courbe p(t) par pas de 250 μs, sur une durée de 500 ms, ce qui correspondait aux performances de l'appareil dont nous disposions. Si l'allure globale de la courbe d'énergie cumulée est :

$$w(t) = 1 - e^{-13,8t/T}$$

où la durée de réverbération T est de l'ordre de 2 s, w(t) atteint 97% de sa valeur maximale au bout de 500 ms.

En estimant que le carré de la pression sonore est proportionnel à l'énergie reçue par le microphone, nous avons obtenu sur ordinateur, les valeurs $w(t_i) \simeq p^2(t_i)$ qui permettent de tracer la courbe d'énergie cumulée :

$$w(t_n) = \sum_{i=1}^{n} \Delta w(t_i)$$

d) Mesure du coefficient de réflexion des parois.-

Lorsqu'on veut comparer le calcul et l'expérience, il est nécessaire d'introduire dans le calcul des coefficients de réflexion des parois dont la valeur ait quelque rapport avec la réalité. Une méthode est celle de la salle réverbérante dont on mesure la durée de réverbération avant et après l'introduction du panneau absorbant. On en déduit alors une certaine valeur du coefficient d'absorption, en utilisant une formule de réverbération classique (Sabine, Eyring, Millington). Une autre méthode, que nous avons été amenés à utiliser, est la méthode "d'incidence", d'une onde plane. Cette méthode est délicate à mettre en oeuvre, mais on peut estimer que c'est la seule utilisable pour une mesure "in situ". Dans nos calculs, nous avons utilisé seulement un coefficient d'absorption moyen, mais il est également possible d'introduire un coefficient variable avec l'angle d'incidence.

d_1 - Méthode de la salle réverbérante [4.23] [4.24]

Nous reprendrons l'exemple de la salle heptaédrique, utilisée usuellement comme chambre réverbérante. On peut supposer, que lorsque cette salle est nue, toutes les parois présentent le même coefficient d'absorption. Les mesures de réverbération ont été effectuées en utilisant une bande d'une octave de largeur, centrée sur 1000 Hz. Dans ces conditions, on a

mesuré une durée de réverbération T_0 située entre 5,6 et 7 secondes. On en déduit une valeur du coefficient d'absorption des murs

$$\alpha_o = 0,16V/ T_0S$$

où V est le volume de la salle (V = 173 m^3) et S sa surface totale (S = 179m^2) α_0 est compris entre 0,028 et 0,022, ce qui donne un coefficient de réflexion de l'ordre de 0,97 à 0,98.

L'introduction d'un panneau absorbant de surface s, disposé sur un des côtés de la salle fait tomber la durée de réverbération T_0 à une valeur T_1 comprise entre 2,3 et 2,8 s. Dans ces conditions, on obtient une valeur du coefficient d'absorption α_1 du panneau par la relation :

$$T_1 = \frac{0,16\ V}{\alpha_0(S-s) + \alpha_{1s}}$$

On en déduit une valeur de $R_1 = 1 - \alpha_1$ de l'ordre de 0,6 à 0,65.

d_2 - Méthode d'incidence [4.25] [4.26]

Le principe de la méthode d'indidence est très simple : on émet une onde plane qui tombe avec un certain angle d'incidence sur l'échantillon dont on veut mesurer le coefficient de réflexion. Dans la direction correspondant au son réfléchi, on place un microphone directif autant que possible et l'on observe le niveau n_1 du son reçu sur un oscilloscope. Au panneau à étudier, on substitue un panneau supposé parfaitement réfléchissant : le niveau du son reçu sur l'oscilloscope est alors $n_0 \geqslant n_1$. On obtient un coefficient de réflexion de la pression sonore sous l'incidence θ, soit :

$|r| = n_1/n_0$ et l'on en déduit le coefficient de réflexion de l'énergie, soit $R = rr^* = |r|^2$.

En chambre sourde, la manipulation se heurte déjà à certaines difficultés pratiques : en effet, l'onde émise n'est jamais rigoureusement plane, si bien que le son direct risque toujours d'atteindre le récepteur, à moins qu'on utilise un véritable tunnel coudé, en disposant l'échantillon dans l'angle. Si l'on veut effectuer une mesure "in situ," il faudra éliminer autant que possible,non seulement le son direct, mais également les réflexions éventuellement dues à d'autres parois. C'est là qu'apparaît l'avantage présenté

par l'utilisation, à l'émission, de trains d'onde de durée limitée, qui permettent de vérifier, soit l'absence d'une onde perturbatrice, soit d'éviter de la prendre en considération, comme le montre l'exemple envisagé par Cops et Myncke [4.26] . Le dispositif expérimental est montré par les figures 4.27 et 4.28 . La chaîne d'émission est constituée par un générateur Schomandl, un découpeur Général Radio, un amplificateur Dynakit et une chambre de compression munie d'un cornet de section rectangulaire. Le dispositif récepteur comprend un micro Sehnheiser très directif et un oscilloscope. Pour obtenir un coefficient d'absorption moyen $\bar{\alpha}$ pour toutes les incidences possibles, on peut mesurer $\alpha(\theta)$ pour diverses valeurs de θ comprises entre 0 et $\pi/2$ et en déduire :

$$\bar{\alpha} = 2 \int_{0}^{\pi/2} \alpha(\theta) \sin\theta . \cos\theta . d\theta$$

Pour un certain nombre de matériaux, M. Vogel avait remarqué que le coefficient d'absorption moyen est proche de celui mesuré sous une incidence de 55° . Dans ces conditions, nous avons utilisé le résultat de la mesure effectuée sous une incidence de 55° . La figure 4.29 donne quelques niveaux proportionnels à la pression sonore, relevés sur l'oscilloscope, en fonction de la distance du trajet acoustique, de l'émetteur au récepteur. Il apparaît bien que l'on n'a pas affaire à une onde plane, et que l'atténuation en fonction de la distance est assez importante. Le coefficient R trouvé varie entre 0,56 et 0,7.

Le mesure du coefficient de réflexion (ou du coefficient d'absorption) par l'une ou l'autre des deux méthodes ne présente pas une grande précision, il faut donc se contenter d'un certain ordre de grandeur qui donnera cependant une indication très utile pour la comparaison avec le résultat des calculs. La précision est d'autant plus faible que le coefficient d'absorption est plus élevé.

e) Salle heptaédrique.-

Nous avions déjà effectué des calculs relatifs à cette salle : nous les avons repris, avec les modifications suivantes : le coefficient de réflexion des parois nues a été évalué expérimentalement à 0,97 : un panneau absorbant a été installé, pour des raisons de commodité, contre un mur, et non sur le sol, où il aurait passablement entravé le déplacement des dispositifs d'émission et de réception. Ce panneau est constitué par des dièdres en mousse

MESURE DU COEFFICIENT
DE REFLEXION D'UN PANNEAU

Fig. 4.27 - Dispositifs d'émission
et de réception -

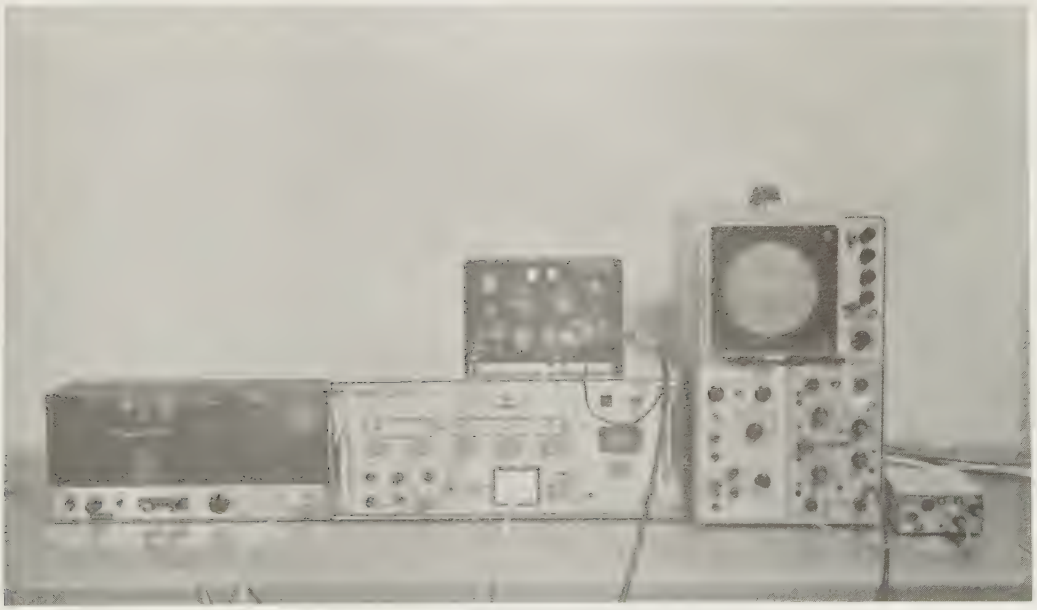
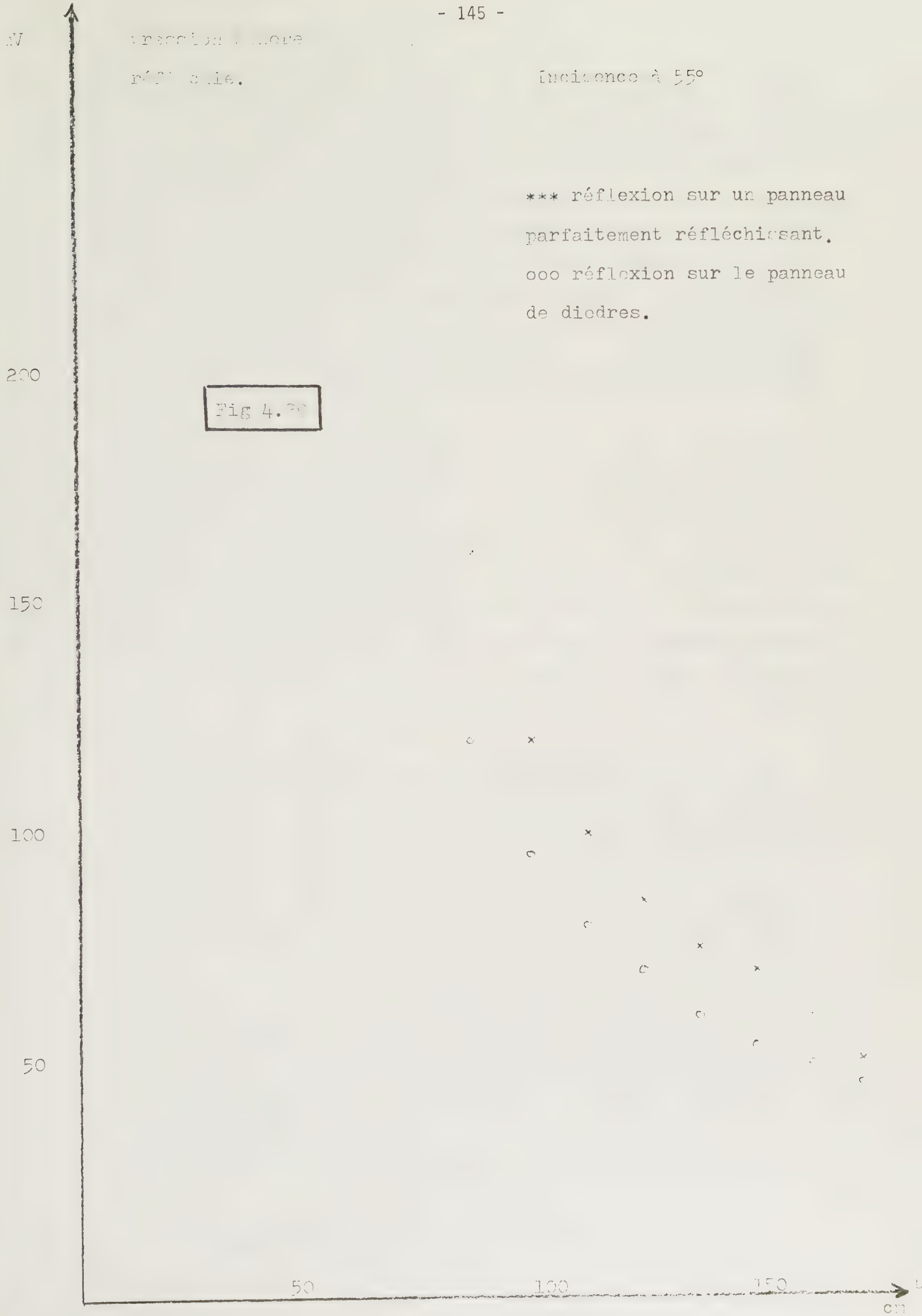


Fig. 4.28 - Alimentation du dispositif d'émission
et oscilloscope de mesure.


Incidence à 55°
*** réflexion sur un panneau parfaitement réfléchissant.
ooo réflexion sur le panneau de diedres.
200
150
100
50
50
100
cm

de polyester, montés sur des plaques de polystyrène expansé (Fig. 4.30). Le coefficient de réflexion de ce panneau est de l'ordre de 0,6 : nous verrons plus loin comment nous avons été amenés à préciser sa valeur. Le calcul a été effectué avec les données suivantes : la source isotrope émet 4000 rayons, le récepteur sphérique a 50 cm de diamètre. Dans les situations envisagées de la source et du récepteur, le récepteur reçoit environ 200 rayons. L'erreur de calcul due au tirage peut être estimée inférieure à 5%. Nous avons déjà présenté le dispositif expérimental utilisé : la source et le récepteur sont visibles sur la Figure 4.31, alors que le dispositif d'émission (Générateur BF, découpeur et amplificateur) et le dispositif de réception (alimentation et enregistreur numérique) sont présentés par la Figure 4.31.

On a envisagé, pour une même situation de la source, trois positions différentes du récepteur. Dans chaque situation, pour comparer la courbe calculée et la courbe expérimentale, on a supposé que les énergies maximales obtenues dans les deux cas étaient identiques.D'autre part, la courbe expérimentale a servi à ajuster la valeur du coefficient de réflexion du panneau absorbant, qui était connu avec une précision médiocre. On a retenu la valeur de 0,6. Les Figures 4.32, 4.33, et 4.34 présentent les résultats obtenus pour les trois positions du récepteur : la concordance s'avère très satisfaisante, les courbes expérimentales entrent dans les limites de l'erreur due au tirage, mentionnée plus haut.

Remarques sur le son direct et le son réverbéré.

Il n'est pas inutile de comparer les résultats de notre calcul numérique avec ceux que l'on pourrait attendre des formules classiques de réverbération, et en particulier de la formule de Sabine.
Pour la commodité du raisonnement, nous supposerons que la source émet à partir de l'instant t = 0 une puissance unité, et non une énergie unité à l'instant t = 0. La courbe w(t) est alors la puissance recueillie au récepteur, elle est identique à la précédente courbe d'énergie cumulée.
Pour estimer la densité de l'énergie réverbérée, nous reprendrons le calcul de Maxfield et Albertsheim [4.27] . Si E_0 est l'énergie totale réverbérée, à un instant donné, doit t = 0, supposons que l'on arrête la source : l'énergie décroît selon la loi :

$$E = E_0 \, e^{-13,8t/T}$$

où T est la durée de réverbération.

ENREGISTREMENT D'UN ECHOGRAMME

Fig. 4.30 - Salle heptaédrique et panneau absorbant -

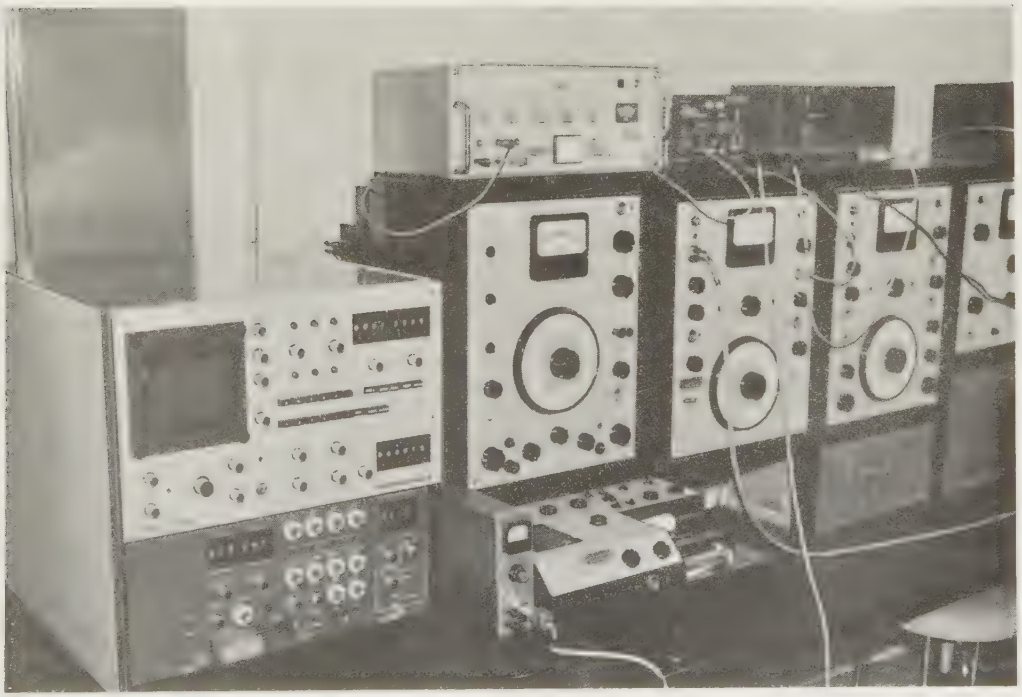


Fig. 4.31 - Dispositifs d'émission et de réception -

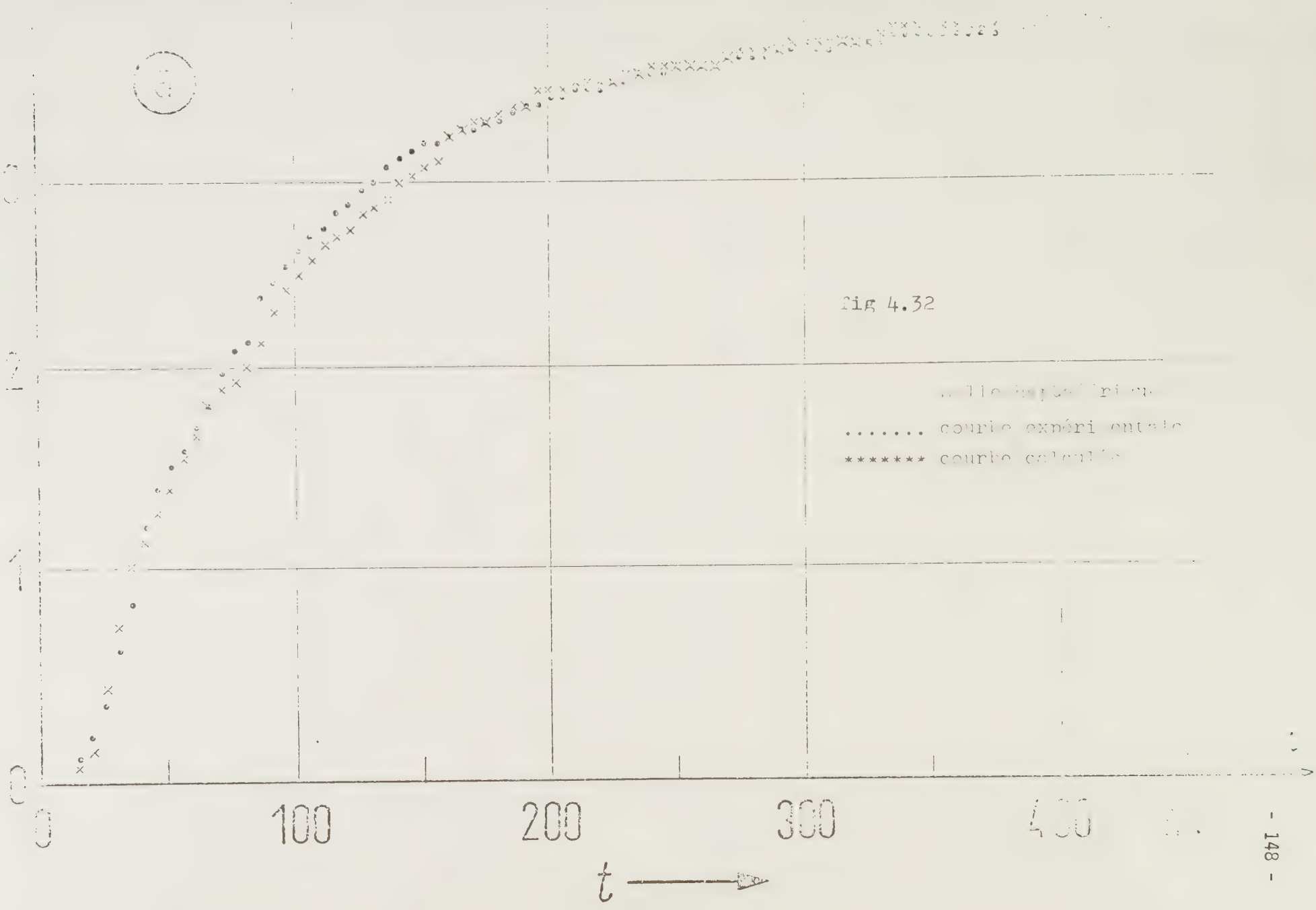


Fig 4.32

..... courbe expérimentale

******* courbe calculée

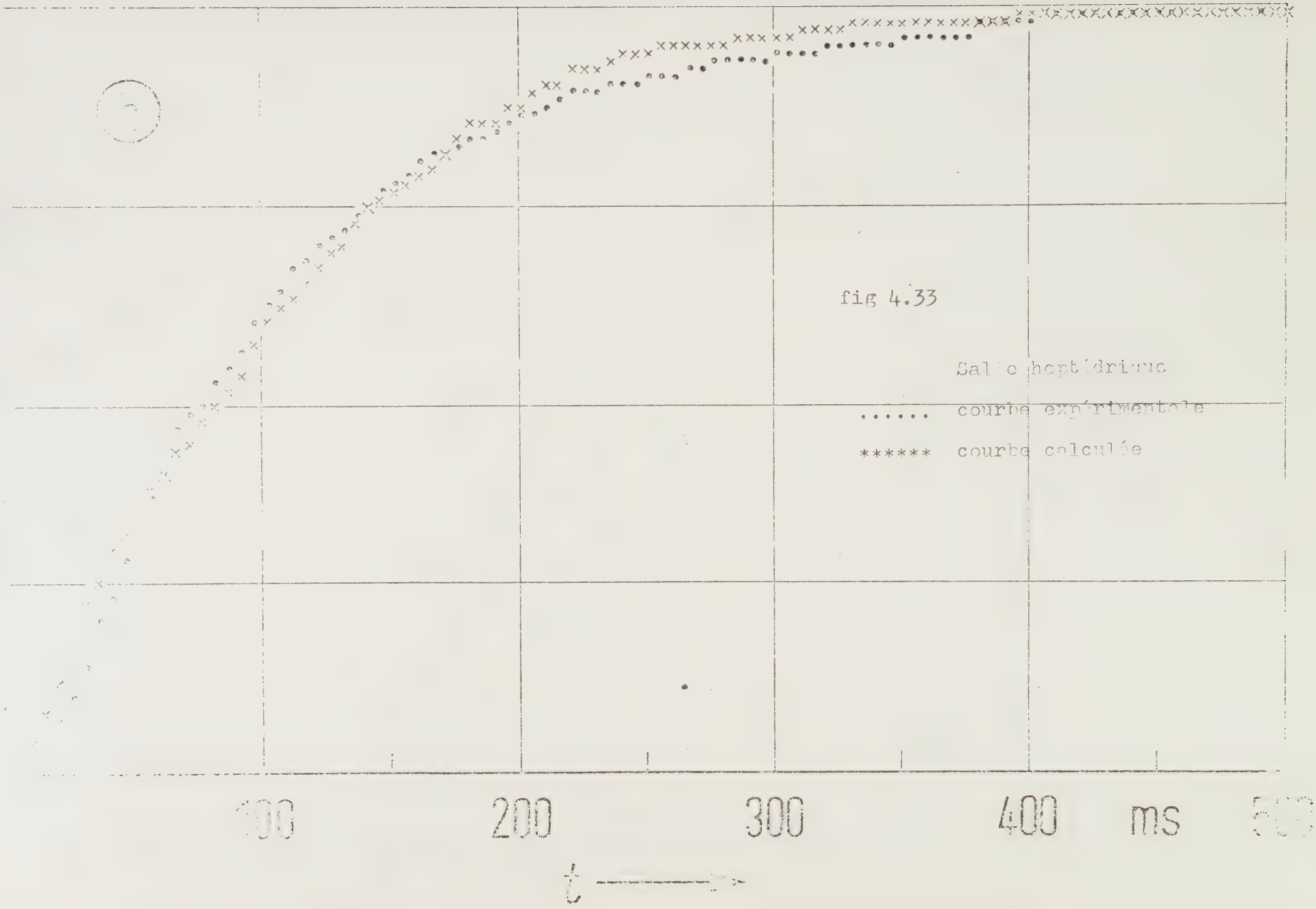


fig 4.33

Sal'c hept'drique

..... courbe expérimentale

****** courbe calculée

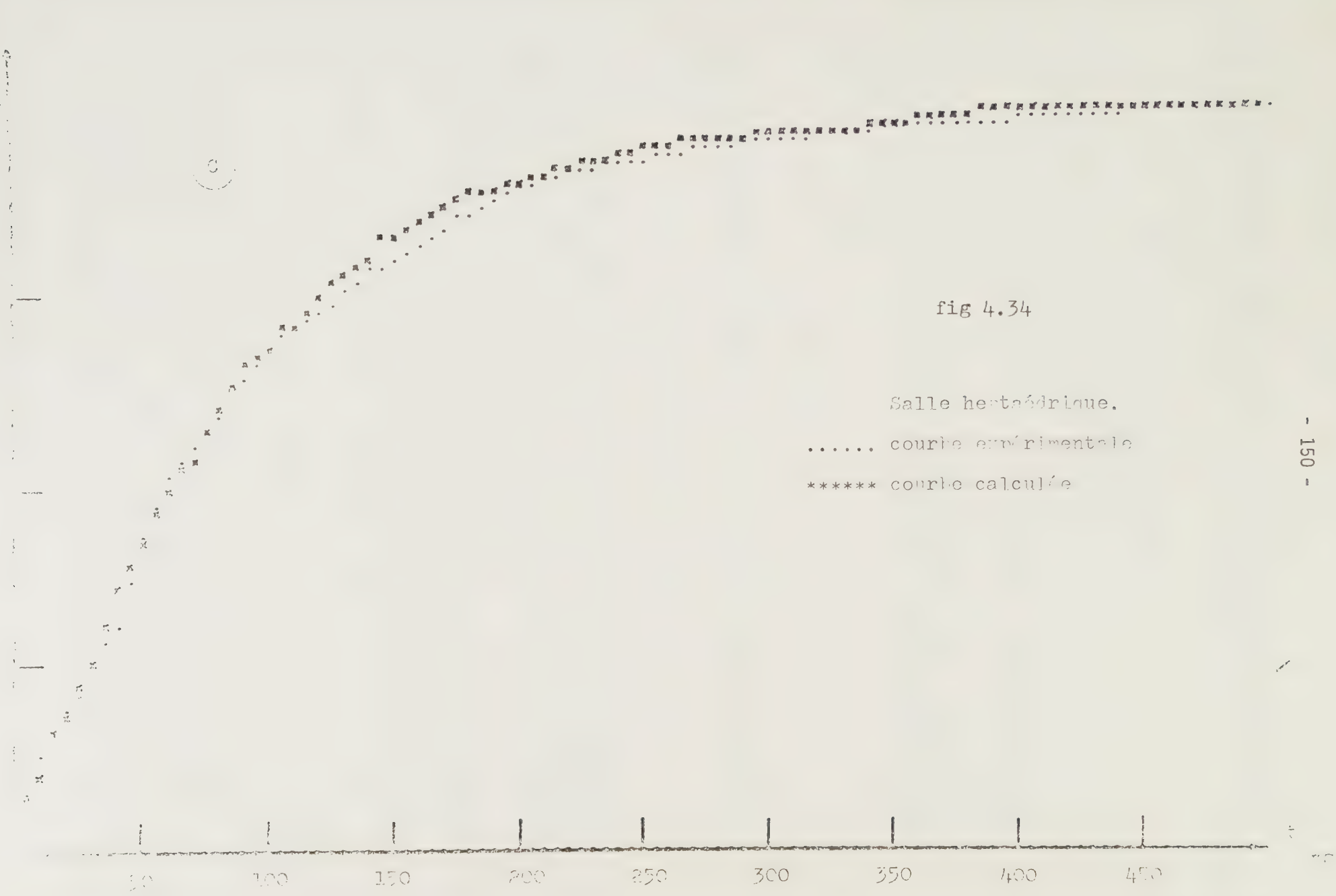


fig 4.34

Salle hexaédrique.

..... courbe expérimentale

****** courbe calculée

En régime permanent, lorsque la source fonctionne, on peut estimer que la puissance w fournie par la source compense l'absorption du son par les parois, et que, par conséquent

$$w = dE/dt = 13{,}8\ E_o/T$$

La densité d'énergie réverbérée est alors :

$$e_n = E_o/V = WT /\ 13{,}8\ V\ ,\ \text{si V est le volume de la salle.}$$

La puissance réverbérée reçue est l'énergie contenue dans un cylindre de longueur c, célérité du son, et de section πr^2 , surface apparente du récepteur sphérique :

$$w_r = WT.\ c\pi r^2 / 13{,}8\ V$$

Si D est la distance du récepteur à l'émetteur, la puissance du son direct sera :

$$w_d = W.\pi r^2 / 4\pi D^2$$

Pour la première situation envisagée, repérée a, la puissance du son direct reçu, soit w_d varie entre $0{,}05.10^{-2}$ et $0{,}1.10^{-2}$ suivant les tirages : on suppose que la source émet une puissance unitaire. La puissance w_d calculée suivant l'expression ci-dessus est de $0{,}078.10^{-2}$.
Il ne faut pas s'étonner de cette bonne concordance, puisque le calcul numérique donne une valeur de w_d directement associée à cette expression.

Par contre, les résultats obtenus pour w_n relèvent de schémas différents. Par l'expression littérale précédente, on obtient :

$$w_n = 0{,}05 \quad \text{en prenant} \quad T = 1{,}8\ s$$

alors que les divers résultats obtenus pour les situations a, b, et c donnent une valeur de l'ordre de 0,04 pour w_r . Rappelons que le résultat de l'expression ci-dessus est obtenu en supposant une équipartition de l'énergie dans tout l'espace.
Il apparaît que le calcul numérique confirme, tout au moins dans le cas de

cette salle, l'ordre de grandeur du résultat obtenu par un calcul de réverbération.

f) Salle de conférences du CNRS à Marseille.-

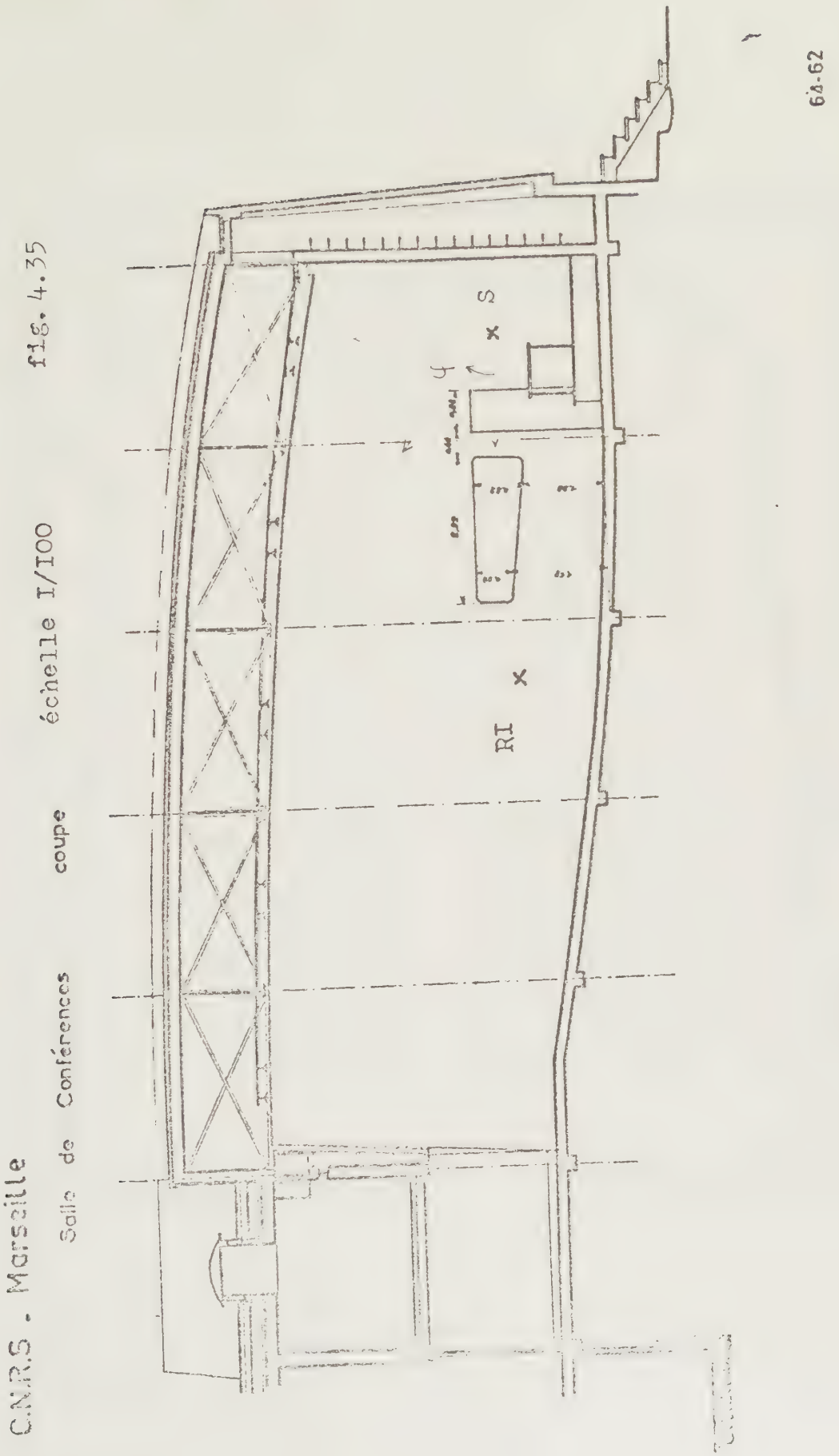
La salle de conférences du Groupe des Laboratoires du CNRS à Marseille est une salle de 176 places. Ses dimensions et sa géométrie sont indiquées par les figures 4.35 et 4.36. Elle est caractérisée par des murs verticaux pratiquement nus : derrière la chaire, un mur incurvé est raccordé aux deux murs latéraux par deux portions de cylindre. Le panneau du fond est constitué par une sorte de grillage en bois derrière lequel est tendue de la toile de jute. On remarquera le grand rayon de courbure du mur situé derrière le locuteur, ce qui évite des effets de focalisation toujours fâcheux, et d'autre part la présence de deux murs latéraux non parallèles, ceci pour empêcher la présence d'échos flottants.

Pour le calcul numérique, cette salle a été considérée comme délimitée par trois plans (panneau du fond, murs latéraux) et cinq portions de cylindres, dont trois d'axe vertical (murs situés derrière le locuteur) et deux d'axe horizontal (sol et plafond). Dans le programme, on a associé à chaque cylindre un plan défini par deux génératrices correspondant aux limites des portions de cylindre. Si un rayon rencontre un de ces plans à l'intérieur de la salle, on calcule son intersection avec le cylindre correspondant, sinon, on passe au plan suivant. Bien entendu, la géométrie de la salle fait apparaître certaines zones particulières pour lesquelles on a dû envisager un traitement numérique spécial.

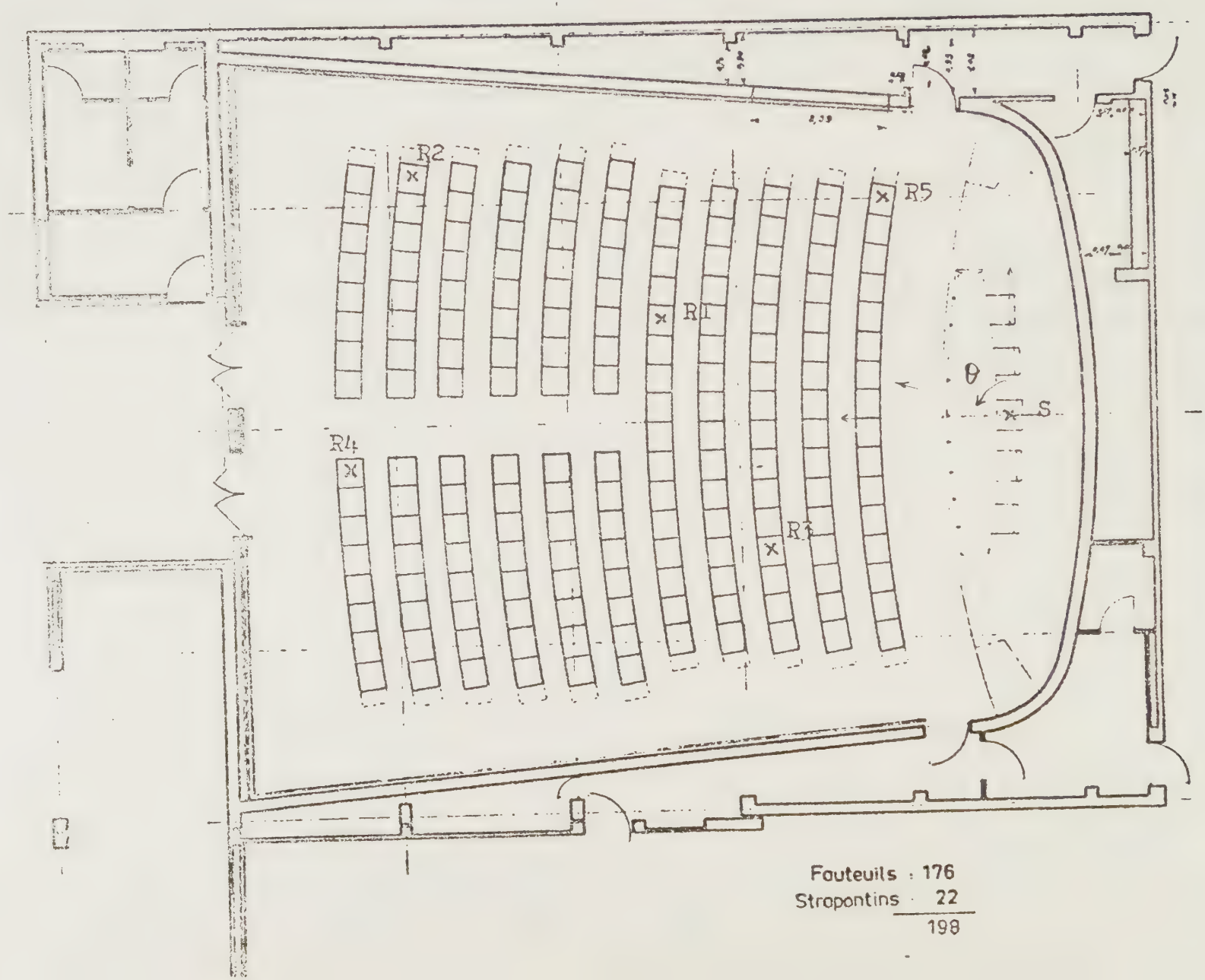
Le coefficient de réflexion moyen des parois a été évalué par une méthode d'incidente à 0,65 pour le panneau du fond et à 0,98 pour les murs et la zone réfléchissante du sol : le reste, occupé par des fauteuils est supposé parfaitement absorbant.

Le rayonnement de la source isotrope a été discrétisé en 2000 rayons, et le récepteur assimilé à une sphère de 2m de diamètre. Ces deux valeurs ont été retenues, pour deux raisons principales.

- Le récepteur reçoit suffisamment de rayons (entre 150 et 200 suivant la place envisagée) pour que les résultats ne diffèrent guère suivant la répartition aléatoire des rayons pris à l'émission : la figure 4.37 présente deux courbes w(t) obtenues pour deux tirages différents des rayons sonores, le récepteur occupant la place R1 indiquée sur la figure 4.35.


C.N.R.S. - Marseille
Salle de Conférences
coupe
échelle I/100
fig. 4.35
64.62
S
RI
4

CNRS - Marseille Salle de Conférences . Plan - échelle I/100 fig. 4.36



Fauteuils : 176
Strapontins : 22
198

64-61

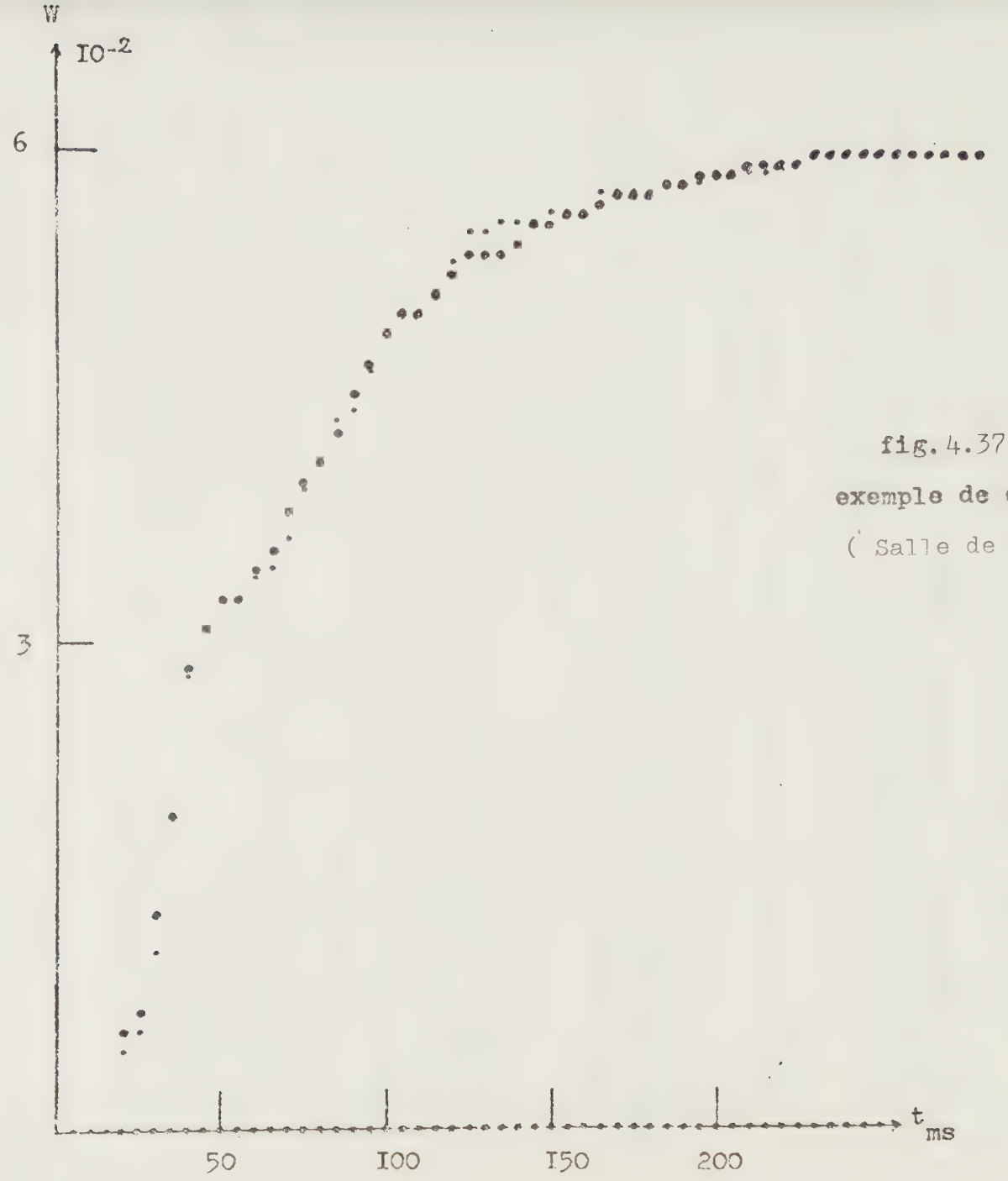


fig. 4.37

exemple de deux tirages

(Salle de conférences)

- L'étude de l'histoire de 2000 rayons dans la salle considérée est exécutée en moins de 2 minutes sur ordinateur Univac 1106. Les calculs devaient être répétés assez souvent, et il était par conséquent avantageux de choisir une dépense en temps de calcul relativement limitée. Les figures numérotées de 4.38 à 4.42 montrent les courbes w(t) obtenues dans ces conditions lorsque le récepteur occupe les places R1, R2, R3, R4 et R5 respectivement. La source garde la situation S indiquée sur les figures 4.35 et 4.36.

Les figures 4.43 et 4.44 donnent deux exemples d'échogrammes spatio-temporels calculés, tout au moins en leur début pour les situations R1 et R2. Les rayons ont des coordonnées angulaires θ et φ telles qu'elles sont indiquées sur les figures 4.35 et 4.36.

A titre d'exemple, nous donnons les résultats de deux calculs effectués avec les hypothèses suivantes :

- On suppose la source anisotrope, ce qui est le cas d'un locuteur réel, et présentant un affaiblissement de 8 dB vers l'arrière : la courbe de la figure 4.45 montre l'énergie cumulée en R1, dans ces conditions. On observe dans ce cas, une augmentation de la netteté.

- On suppose que le panneau du fond est parfaitement diffusant : un rayon s'y réfléchit suivant une direction prise au hasard dans le demi-espace de la normale intérieure à la salle. La figure 4.46 montre le résultat obtenu. Dans ce cas également, on peut constater une augmentation de la netteté par rapport au cas d'un panneau présentant une réflexion spéculaire.

La vérification expérimentale a été effectuée de façon analogue au cas de la salle heptaédrique. Les figures 4.38 et 4.42 montrent simultanément les résultats expérimentaux et ceux du calcul. On a porté sur le tableau ci-dessous l'écart maximum entre la courbe calculée et la courbe expérimentale par rapport à la valeur de l'énergie totale cumulée, pour chaque situation envisagée.

Place du récepteur	R1	R2	R3	R4	R5
Ecart maximum %	6	14	9	9	12

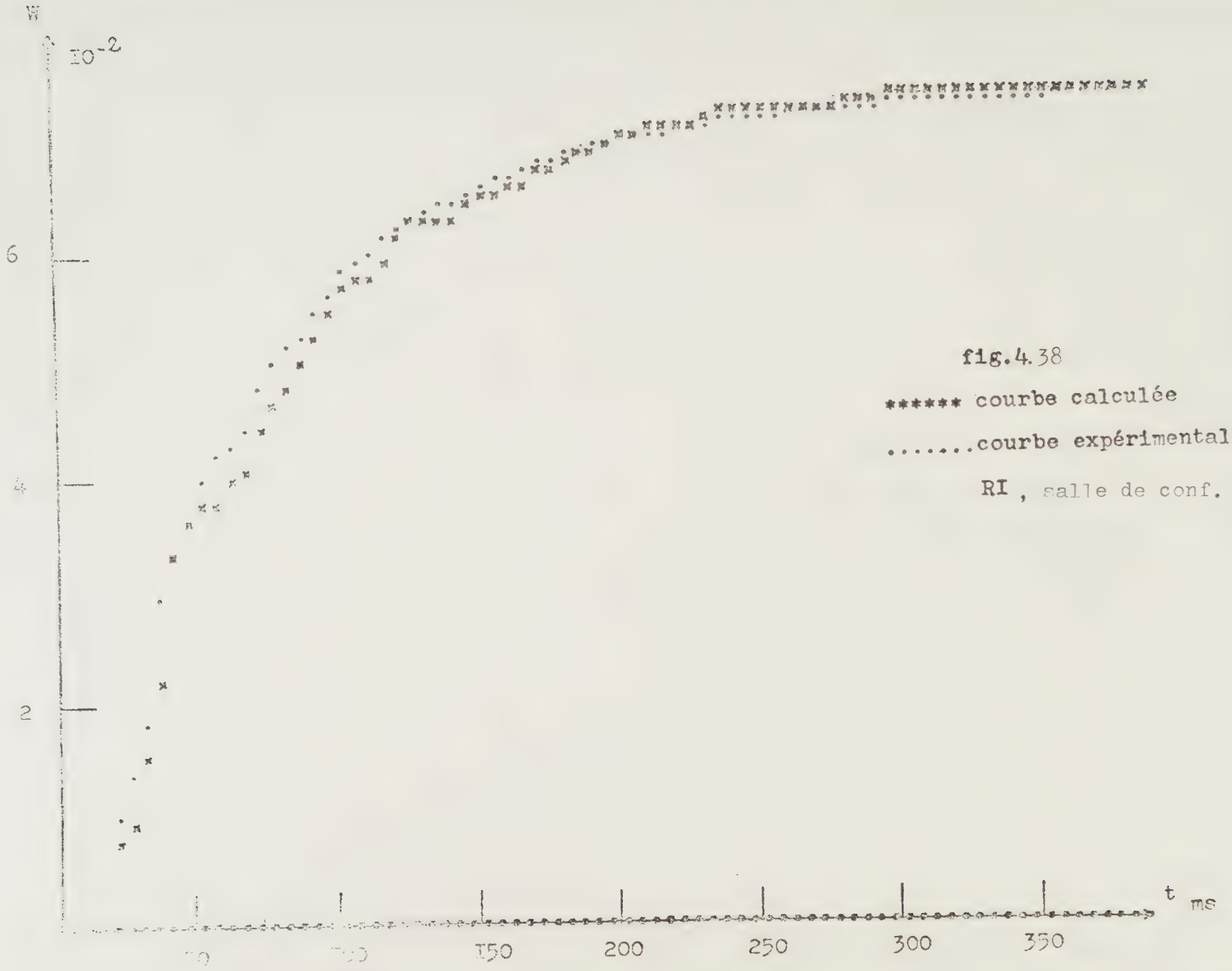


fig.4.38

****** courbe calculée

.....courbe expérimentale

RI , salle de conf.

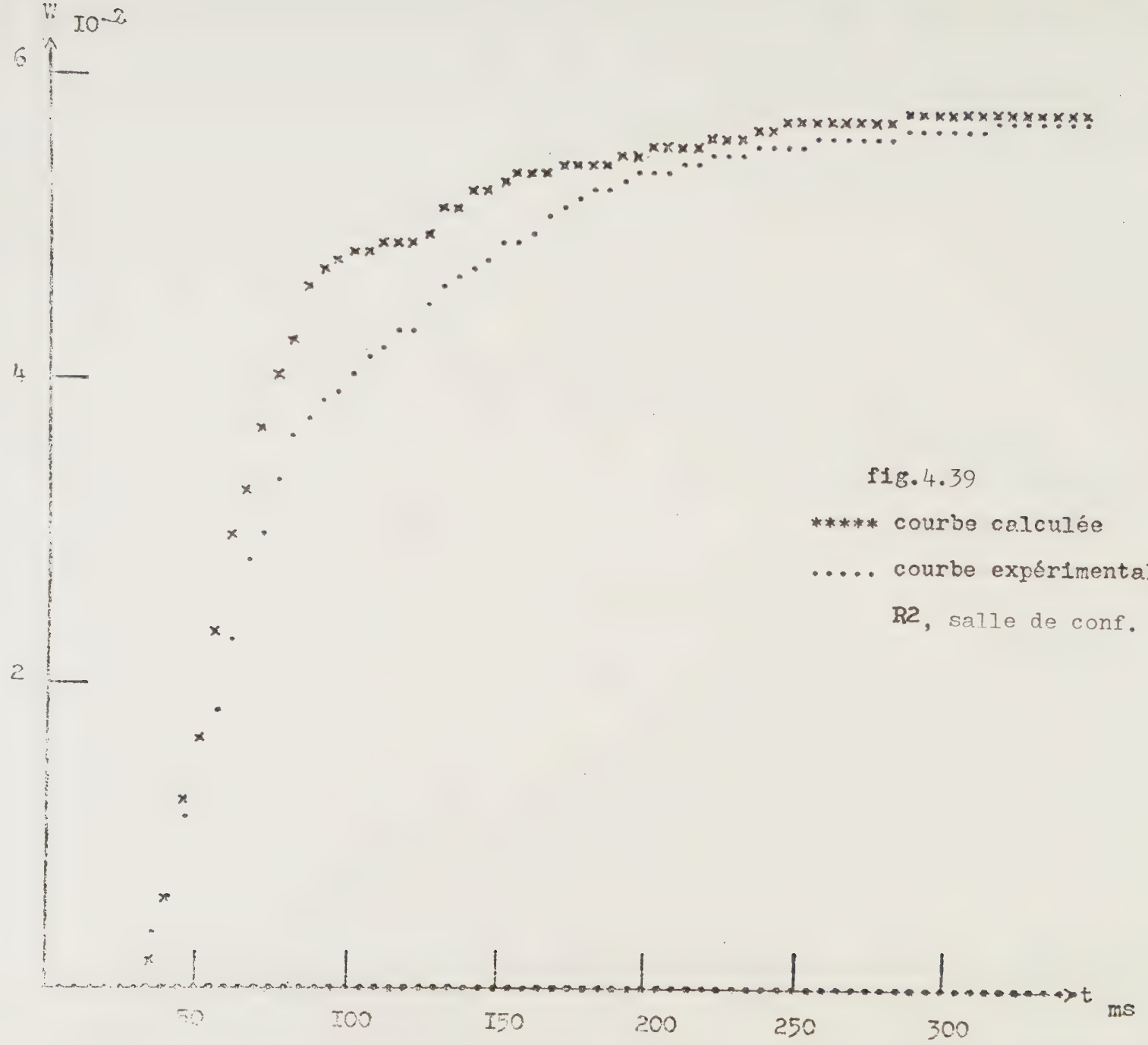


fig.4.39

***** courbe calculée

..... courbe expérimentale

R2, salle de conf.

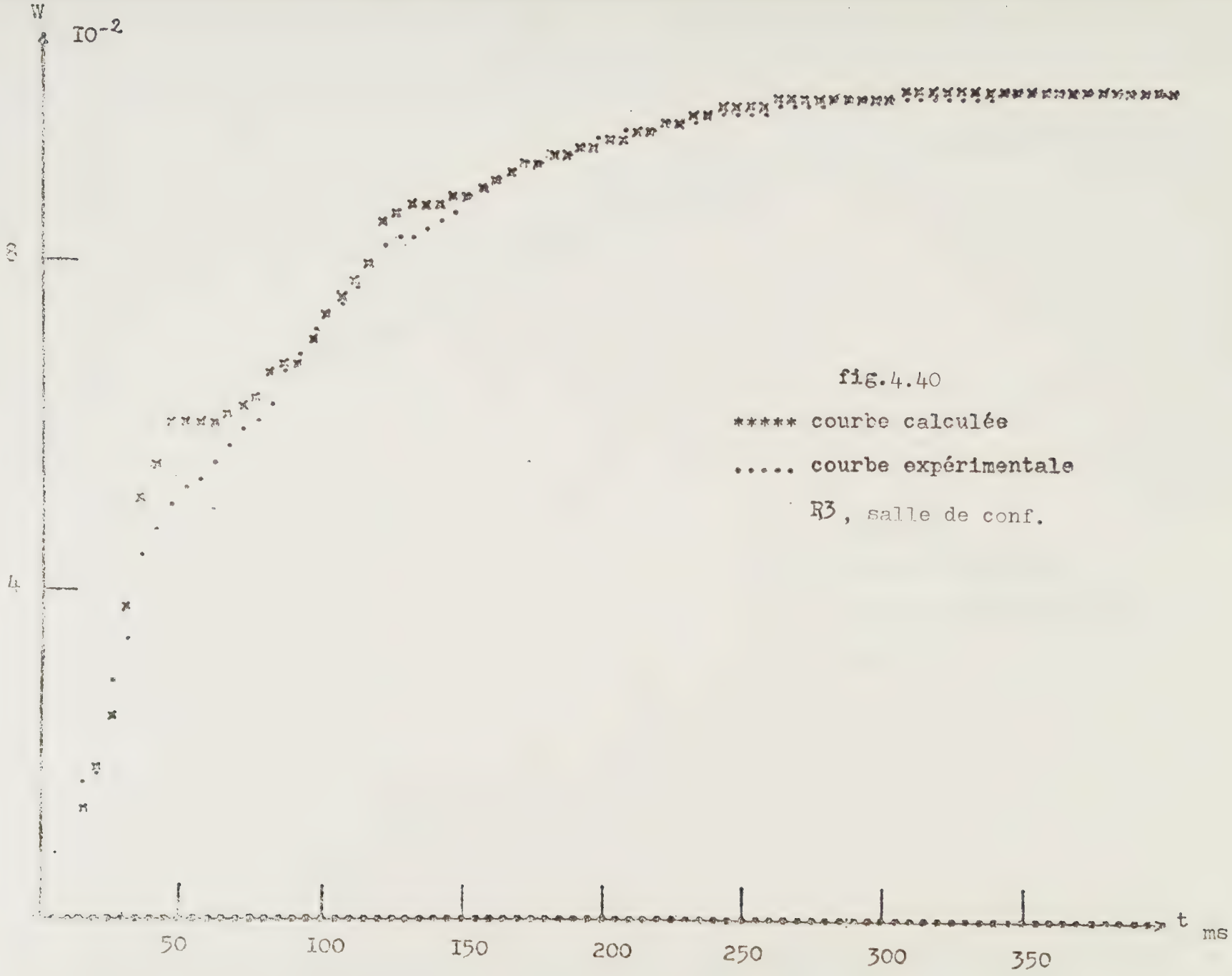


fig.4.40

***** courbe calculée

..... courbe expérimentale

R3, salle de conf.

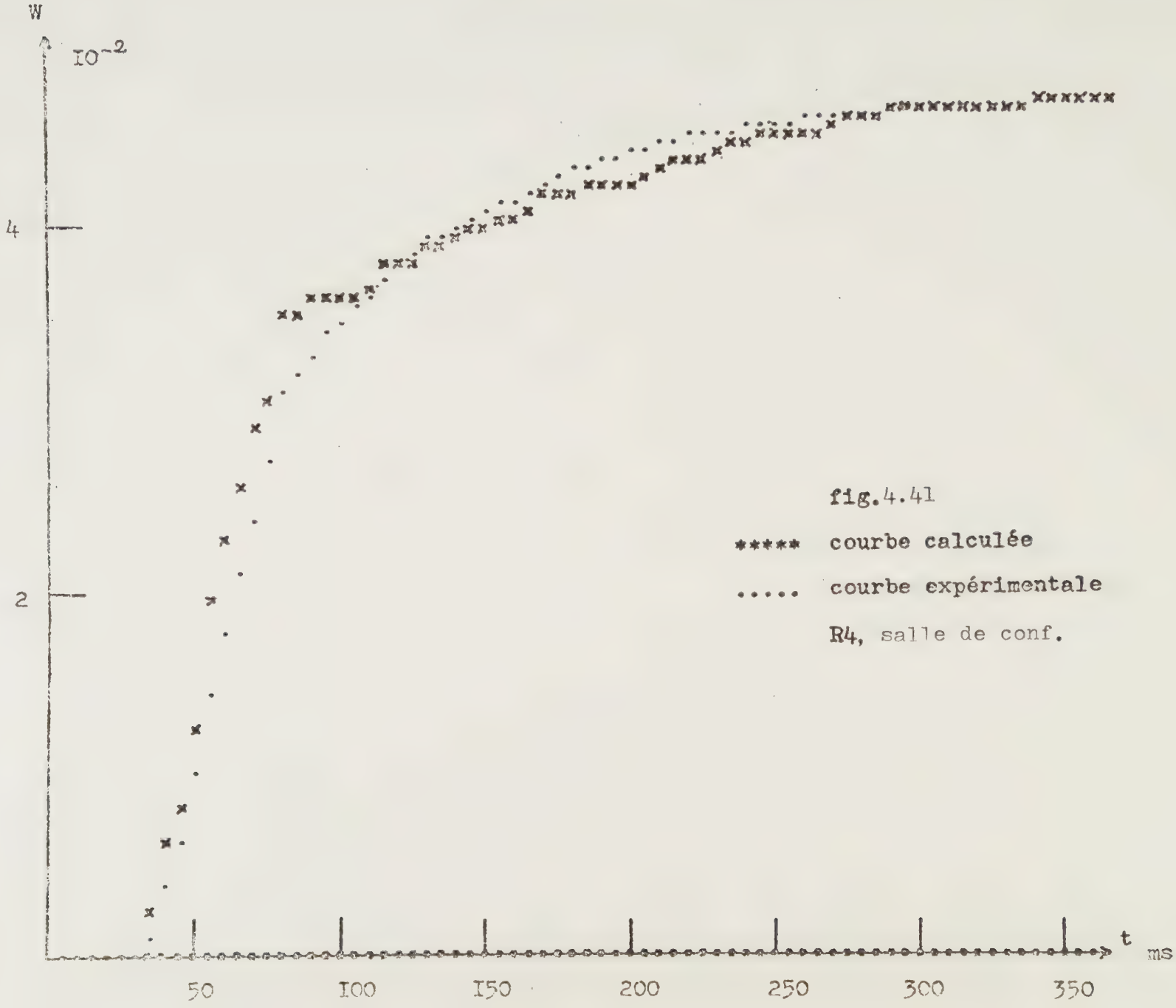


fig.4.41

***** courbe calculée

..... courbe expérimentale

R4, salle de conf.

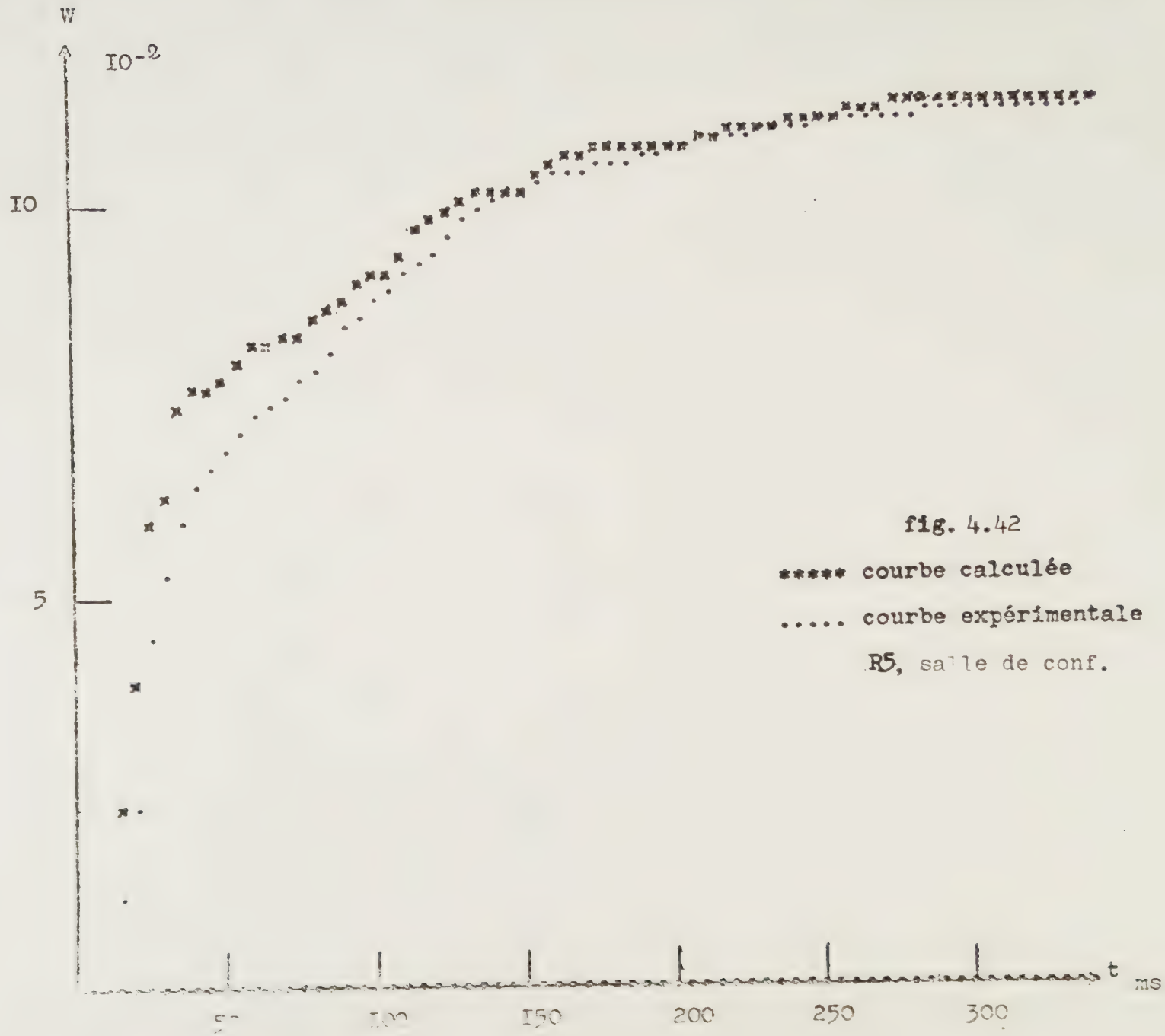


fig. 4.42

***** courbe calculée

..... courbe expérimentale

R5, salle de conf.

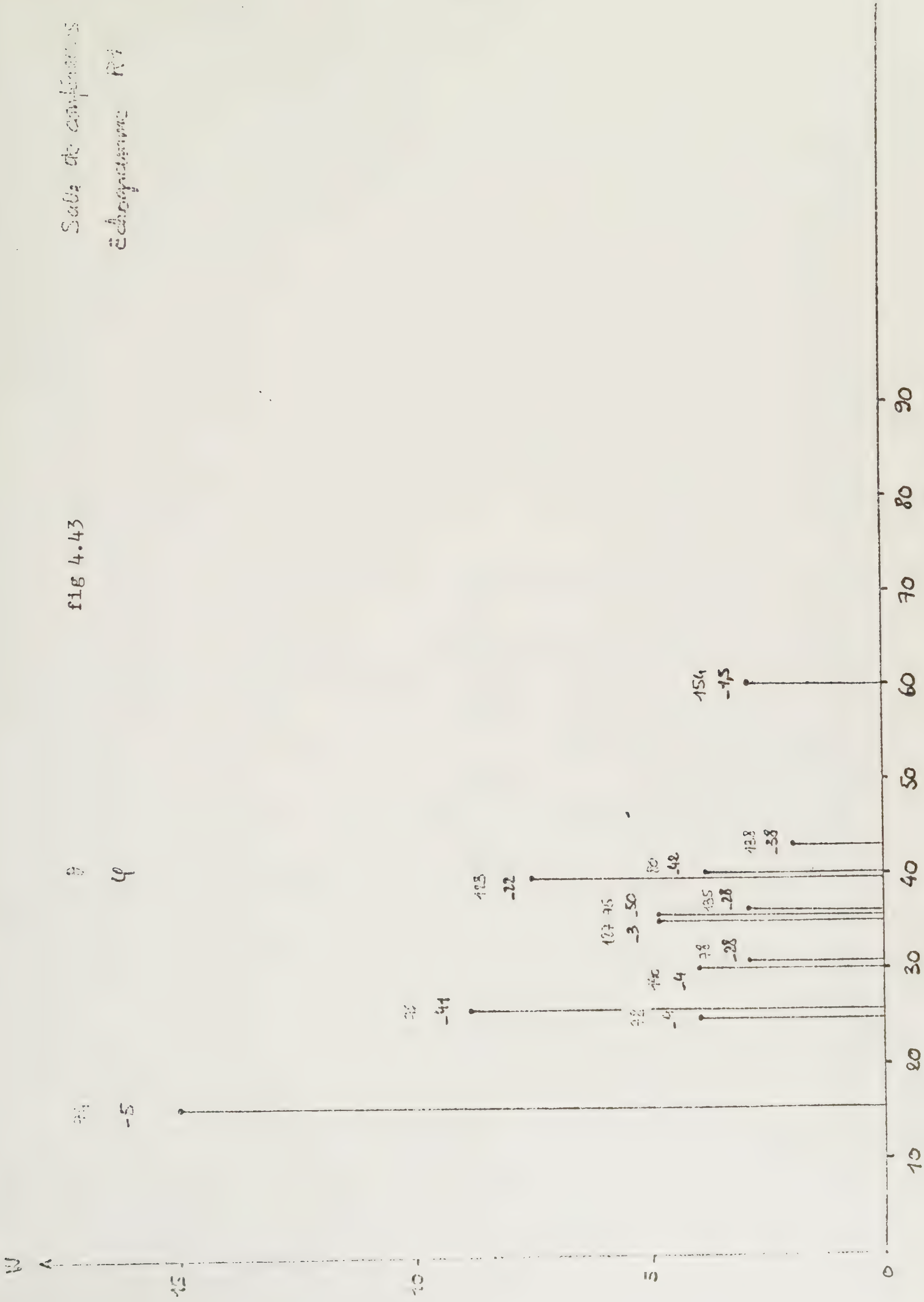
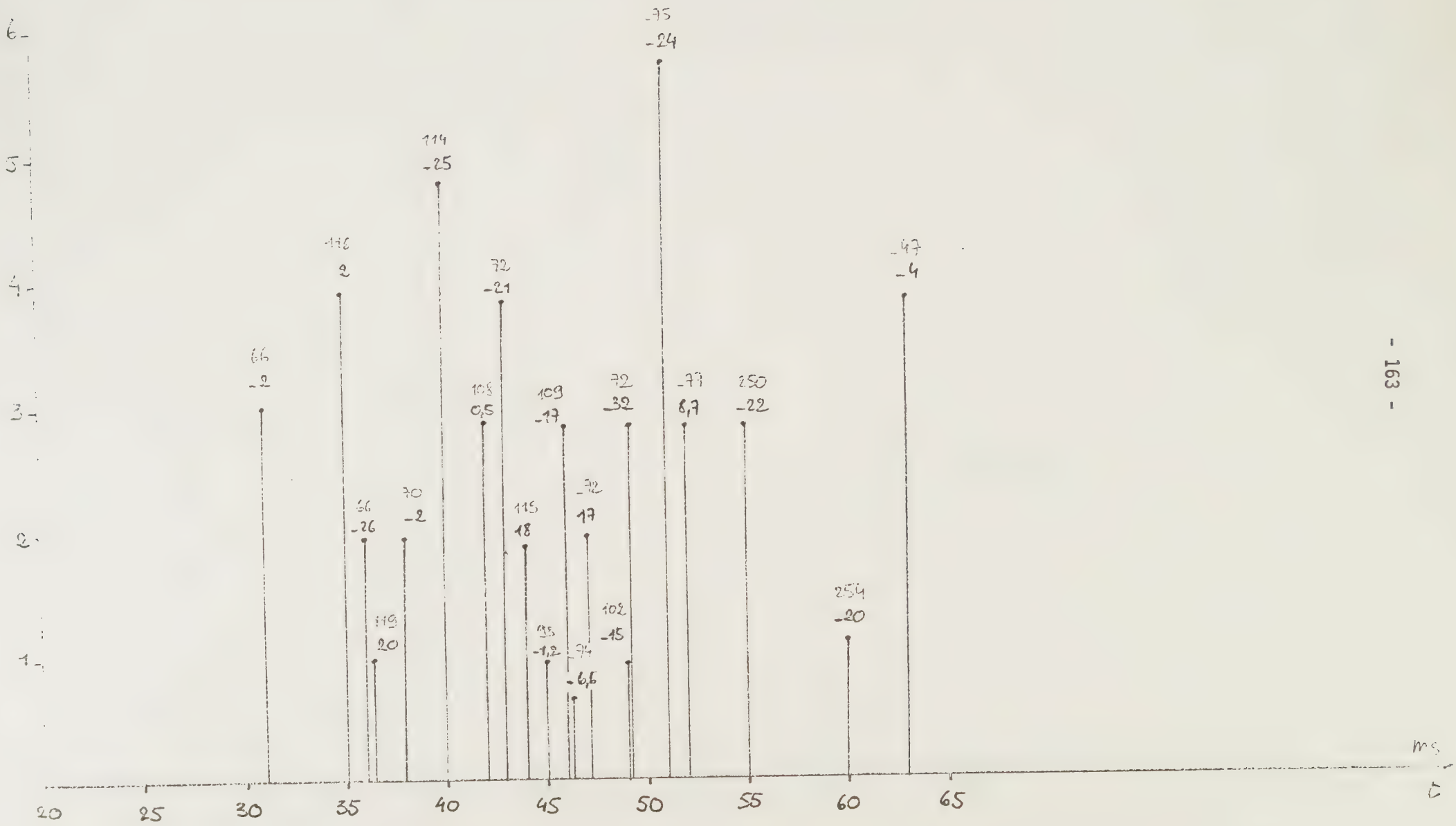


fig 4.43

fig 4.44


6
5
4
3
2
1
20
25
30
35
40
45
50
55
60
65
ms
66
-2
116
2
66
-26
119
20
70
-2
114
-25
108
0,5
72
-21
115
18
109
-17
95
-1,2
94
-6,5
92
17
72
-32
102
-15
75
-24
73
8,7
150
-22
254
-20
47
-4

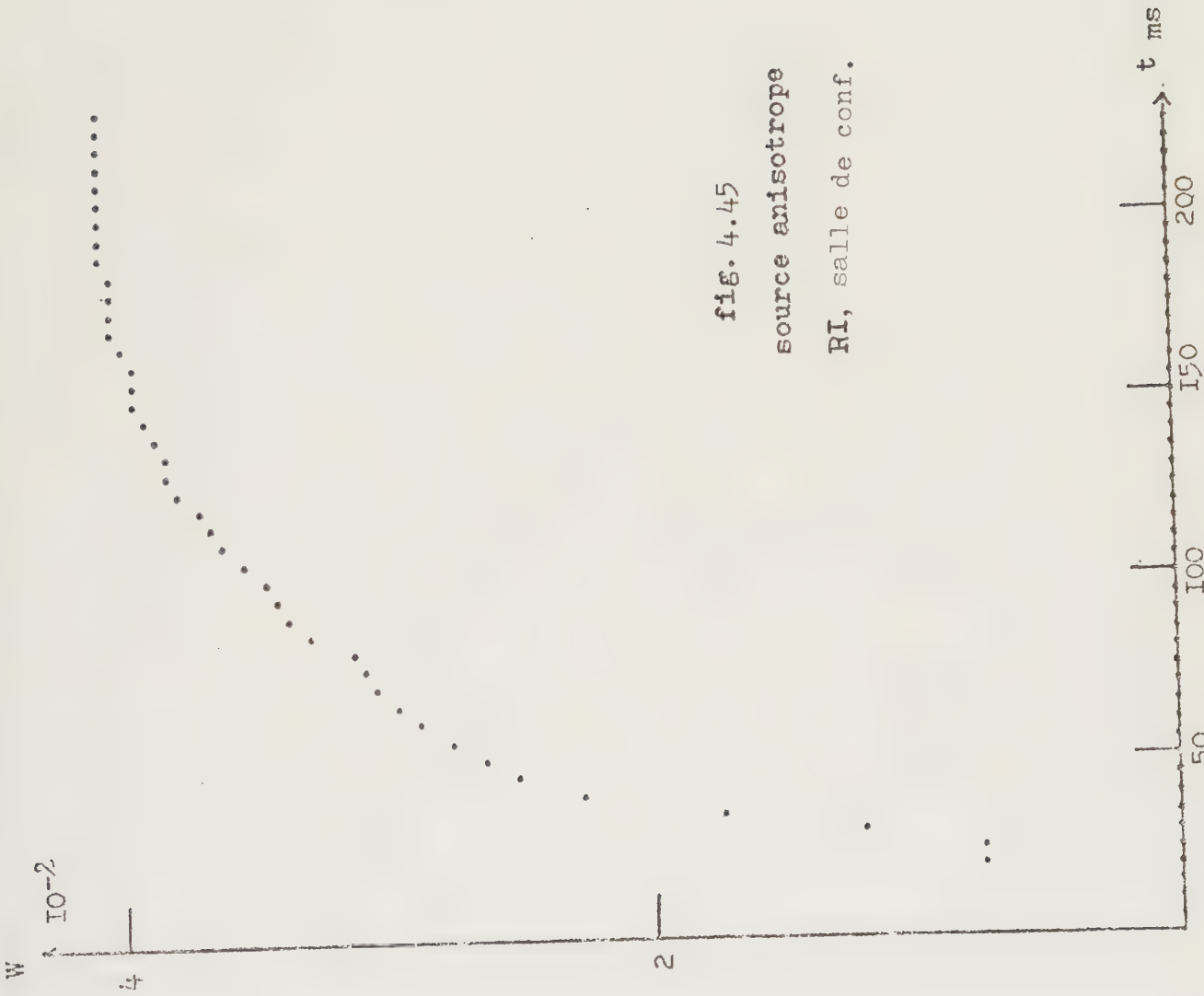


fig. 4.45
source anisotrope
RI, salle de conf.

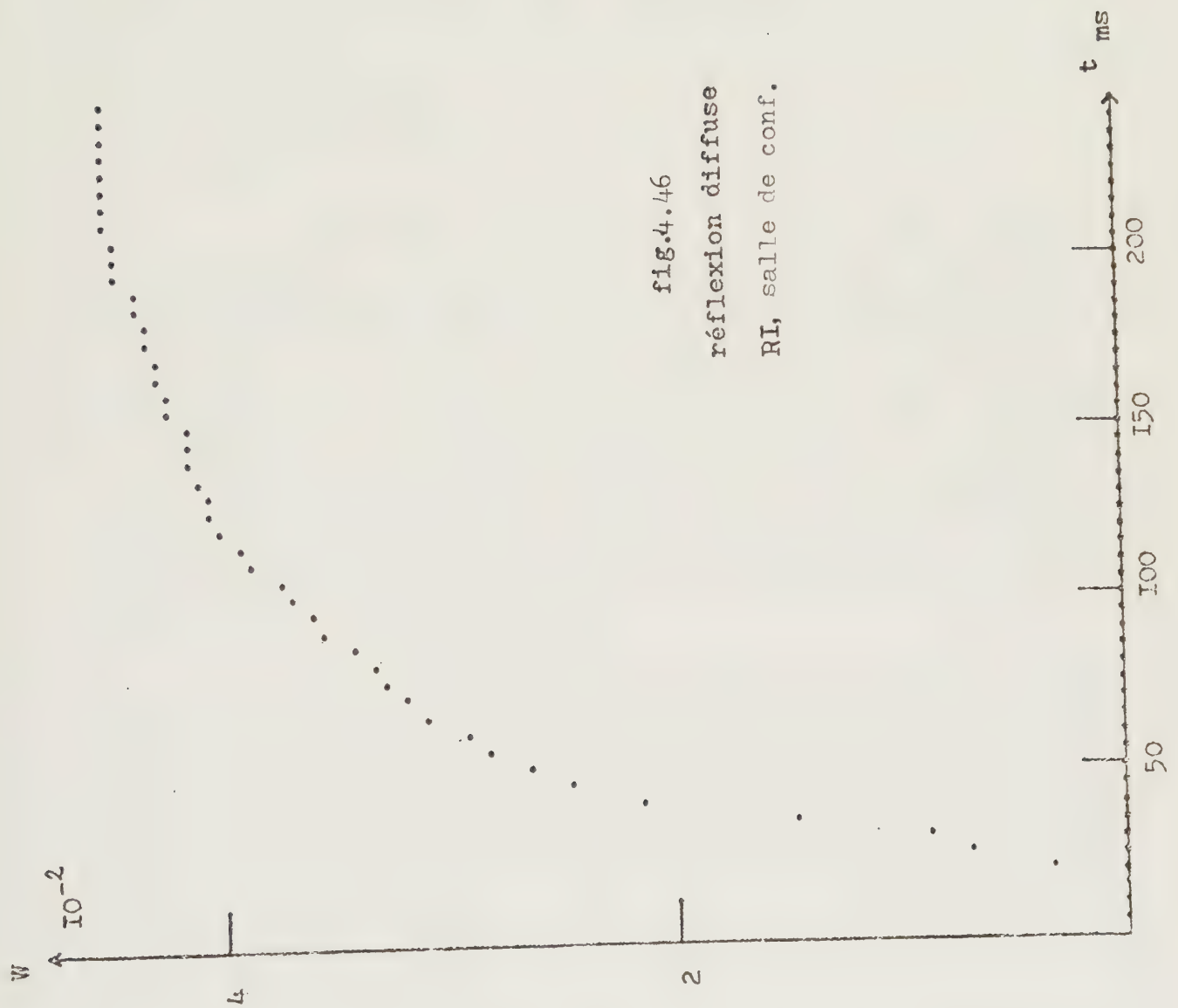


fig.4.46
réflexion diffuse
RI, salle de conf.

On peut estimer que la concordance est bonne, si l'on considère d'une part l'imperfection de l'impulsion expérimentale, d'autre part, les simplifications faites sur la forme de la salle, pour ne pas trop alourdir les calculs: on a négligé la présence d'une estrade, d'une table de projection, l'épaisseur de la zone absorbante constituée par les fauteuils et les décrochements pratiqués dans le panneau du fond, correspondant aux portes, etc...

4.4.- Conclusion

Au terme de ce chapitre, il importe d'en tirer les points essentiels. Dans une salle de forme pratiquement quelconque, nous avons pu calculer un échogramme spatio-temporel. La vérification expérimentale a été faite à partir de la courbe d'énergie cumulée et donne des résultats très encourageants. Il s'agit donc d'une technique de calcul dont l'efficacité est appréciable, le seul problème est la mise au point d'un programme qui sera d'autant plus compliqué que la forme de la salle envisagée sera elle-même plus complexe. Mais on peut s'estimer dispensé de construire une maquette. Il sera alors possible de connaître l'échogramme dans une salle, en considérant uniquement les plans.

Certains pourraient faire remarquer le paradoxe qui existe entre la qualité du résultat obtenu et la grossièreté de la théorie envisagée. S'agit-il d'une "heureuse complaisance des phénomènes"? [4.25] . On peut remarquer à ce sujet que le problème envisagé est relatif à une enceinte fermée, sans angles rentrants, et que la longueur d'onde correspondant à la fréquence centrale du spectre du signal utilisé dans l'expérience est très inférieure aux dimensions de la salle. Ces éléments diminuent l'importance des effets propres à la diffraction.

En utilisant cette technique de calcul, nous ne prétendons pas avoir résolu tous les problèmes que pose l'acoustique des salles. Cependant, il apparaît nettement que ce procédé est suffisant pour envisager le problème particulier de l'intelligibilité, que nous allons maintenant regarder de plus près.

x
x x
x

B I B L I O G R A P H I E

4.1. DUMERY (G.), EXTREMET (P.) et FILIPPI (P.)- Problèmes de diffraction et équations intégrales.- Note CRP n° 1067, Sept. 1968.-

4.2. MORSE (P.M.) et BOLT (R.H.).- Sound waves in rooms.- Rev. Modern. Physics, V 16, n° 2, pp. 69-150, 1944.-

4.3. MORSE (P.M.) et INGARD (U.).- Theoretical Acoustics.- Mac Graw hill, 1970.-

4.4. DOAK (P.E.) et DODD (S.D.).- Some aspects of the theory of diffusion and diffusers.- J. Sound and Vibr., V. 16, n° 1, pp. 89-98, 1971.-

4.5. BARBAT (P.).- Le contrôle de l'acoustique d'une pièce d'habitation par le revêtement des parois.- 7e CIA, Budapest, communication 20 A 21.-

4.6. BERANEK (L.).- Music, Acoustics and Architecture.- J. Wiley and Sons, 1962.-

4.7. EXTREMET(G.P.).- Contribution à l'étude de la propagation du son en enceinte fermée.- Thèse, Fac. des Sc. Aix-Marseille, 1969.-

4.8. EXTREMET (G.P.).- Propagation du son dans une enceinte fermée.- Acustica, V 26, n° 6, pp. 307-314, 1970.-

4.9. EXTREMET (G.P.).- Utilisation du potentiel de simple couche pour l'étude de la répartition du champ sonore dans une salle réverbérante de forme quelconque.- Colloque "Potentiels de couche", Marseille, 2-5 Nov. 1971.-

4.10. CANAC (F.).- Nouvelles recherches sur les théâtres en plein air.- Acustica, V 7, n° 2, pp. 69-74, 1957.-

4.11. CREMER (L.).- Geometrische Raumakustik.- Hirzel Verlag, Stuttgart, 1948.-

4.12. BOLT (R.), DOAK (P.) et WESTERVELT (D.).- Pulse statistics in room acoustics.- JASA, V 22, N 3, pp. 328-340, 1950.-

4.13. HOLTSMARK (J.)- Reverberation as a stochastic process.- 5e CIA, Liège, Sept. 1965, communication H 38.-

4.14 HOLSTMARK (J.).- Reverberation as a stochastic process.- Physica Norvegica, V 12, N 1, pp. 63-69, 1966.-

4.15. PUJOLLE (J.).- Nouveau point de vue sur l'acoustique des salles.- Rev. d'Acoustique, V 18, pp. 21-25, 1972.-

4.16. PUJOLLE (J.).- Nouvelle formule pour la durée de réverbération.- Rev. d'Acoustique, V 19, pp. 107-113, 1972.-

4.17 PUJOLLE (J.).- Nouvelle théorie de la réverbération dans les salles.- Rev. d'Acoustique, V. 27, pp. 263-267, 1973.-

4.18 THO TUONG (N.).- Etude des trajets des ondes sonores dans une salle par la méthode des images à l'aide d'un ordinateur.- Rev. d'Acoustique, V 23, pp. 351-354, 1972.-

4.19 BARRON (M.).- Growth and decay of sourn intensity in rooms according to some formulae of geometric theory.-J. Sound Vib., V 27, n° 2, pp. 186-196, 1973.

4.20 SCHROEDER (M.).- Digital simulation of sound transmission in reverberant spaces.- JASA, V 47, n° 2, pp. 424-431, 1970.-

4.21 KROKSTAD (A.), STRÖM (S.) et SÖRSDAL (S.).- Calculating the acoustical room response by the use of a ray tracing technique.- J. Sound. Vib., V 8, n° 1, pp. 118-125, 1968.-

4.22 KUTTRUFF (H.).- Simulierte Nachhallkürven in Rechteckräumen mit diffusem Schallfeld.- Acustica, V 25, n° 6, pp. 333-342, 1971.-

4.23 ARMAND (J.A.P.).- Application d'une méthode de Monte-Carlo au calcul de l'acoustique géométrique d'une salle.- Projet de fin d'étude, E.N.S.M., Nantes, 1970.-

4.24 BERANEK (L.).- Acoustics Measurements.- Mac Graw Hill, New-York,1967.-

4.25 LAMORAL (R.).- Problèmes d'acoustique des salles et des studios.- Ed. Chiron, Paris, 1970.-

4.26 VOGEL (Th.).- Propriétés acoustiques des matériaux.- L'onde électrique, V 36, n° 350, pp. 428-434, 1956.-

4.27 COPS (A.) et MYNCKE (H.).- Determination of sound absorption coefficients using a tone burst technique.- Acustica, V 29, n° 5, pp. 287-296, 1973.-

4.28 MAXFIELD (J.P.) et ALBERTSHEIM (W.J.)-- An acoustic constant of enclosed spaces correlatable with their apparent liveness. JASA, V 19, n° 1, pp. 71-79, 1947.-

4.29 KORN (T.S.).- Theory of audio information.- Acustica, 22, p. 336, 1969 70.-

X

X X

X

5.- MESURES ET PREVISIONS D'INTELLIGIBILITE.-

Dans ce chapitre, nous traiterons du dernier volet de notre travail : il s'agit de l'utilisation des échogrammes pour la prévision de l'intelligibilité.

Une première partie sera relative à l'utilisation des échogrammes "à plat", c'est-à-dire des simples échogrammes temporels, en faisant abstraction de l'influence de la direction des échos. Nous envisagerons les différentes méthodes proposées à ce jour, ainsi que leur résultat : nous serons amenés à proposer une méthode "hybride" qui utilise plusieurs éléments empruntés à divers procédés. La salle dans laquelle nous avons effectué des mesures d'intelligibilité est la salle de conférences, et nous confronterons ces résultats à ceux, objectifs, des diverses prévisions.

La deuxième partie est plus ambitieuse : il s'agit de savoir quelle influence peut avoir la direction d'une réflexion donnée sur l'intelligibilité. Pour étudier de façon systématique le rôle des échos, nous avons été amenés à utiliser un champ sonore synthétique qui schématise et simplifie le problème. En effet, un échogramme est, à priori, d'une complexité qui peut paraître quasi-inextricable : si l'on veut y voir plus clair, il importe de pouvoir démonter le mécanisme de la perception des échos.

Avant d'envisager les prévisions d'intelligibilité que nous avons effectuées, nous envisagerons d'abord le problème de l'intelligibilité en présence de bruit, et d'autre part, celui des mesures subjectives d'intelligibilité.

5.1.- L'intelligibilité en présence de bruit.-

Parmi les diverses prévisions d'intelligibilité de la parole proposées à ce jour, beaucoup utilisent une formule déterminée empiriquement : pour certains auteurs, l'intelligibilité suit une certaine loi en fonction de la durée de réverbération, de la netteté, du degré d'écho ou encore du centre de gravité de l'échogramme.

Il nous est apparu plus satisfaisant de nous reporter, comme le fait Lochner [3.30] [3.31] , à un réseau de courbes tracées au préalable, et qui donnent l'intelligibilité en fonction du niveau de la parole, en prenant un bruit de masque comme paramètre : ce bruit a le même spectre que la parole.

En effet, nous considérons que l'intelligibilité de la parole dans une salle dépend du niveau absolu de deux valeurs : l'énergie utile d'une part, ou l'énergie associée au son direct, et l'énergie nuisible d'autre part, celle des échos qui ne sauraient être associés ou intégrés au son direct. Nous aurons l'occasion de revenir sur ces divers problèmes.
Certains objecteront que le brouillage apporté par les échos n'est pas de même nature qu'un bruit, même si ce bruit a le spectre moyen de la parole. Rien en fait ne permet d'affirmer que ce brouillage est différent. Un bruit est, dans la pratique, un signal nuisible, même si pour le physicien, il a une structure plus précise : dans une salle de restaurant, la conversation qui se déroule à une table voisine, si elle ne présente pour moi aucun intérêt, et vient troubler les propos échangés avec mes convives, sera bel et bien un bruit, même si son spectre instantané n'est pas celui d'un bruit.
En tout cas, quelque soit l'idée qu'on se fasse à priori des effets de ces deux types de brouillages, les résultats de Lochner, ou les nôtres montrent qu'il s'agit bien, en définitive, de la même chose.

Un autre problème se pose, qui n'est pas négligeable : l'énergie nuisible correspond à des échos assez lointains, donc à des réflexions qui peuvent avoir, de par le rôle du coefficient d'absorption des parois, variable avec la fréquence, un spectre différent de celui associé à l'énergie utile. On aura alors affaire à l'intelligibilité de la parole en fonction d'un bruit de spectre légèrement différent. Nous ne nous attacherons pas à ce problème que nous considérerons comme connu, de par les nombreuses publications ayant pour objet ce cas, depuis les travaux de French et Steinberg [3.15] jusqu'à ceux de Lehmann [3.10] et de Kryter [5.1] [5.2] .
Il ne faut pas perdre de vue que l'intelligibilité de la parole est une fraction aux dimensions restreintes du vaste domaine de l'intelligibilité en présence de bruit, ou de l'intelligibilité dans les télécommunications. C'est pourquoi nous ne nous arrêterons pas outre mesure sur certains problèmes spécifiques de l'intelligibilité en milieu bruyant. Notre but est, essentiellement, l'analyse des répercussions de l'acoustique d'une salle sur l'intelligibilité. Nous ne nous attarderons pas sur les questions spécifiques d'intelligibilité.

Nous avons déjà indiqué, au paragraphe 3.1, quel était le matériau utilisé pour les mesures d'intelligibilité : il s'agit de 20 listes de 50 logatomes, du type consonne-voyelle-consonne. Chaque liste est précédée par un son de référence à 1000 Hz, qui sert à fixer le niveau de la parole. Ce son est situé à 6 dB au-dessous des maximums de la parole et l'on peut estimer qu'il a la même sonie. Dans ce domaine également, on a proposé certaines méthodes pour fixer le niveau de la parole de façon moins empirique [5.3] [5.4] . Nous ne nions certes pas l'intérêt de cette évaluation, mais elle ne relève pas directement de nos préoccupations, et nous nous contenterons d'une estimation plus sommaire.

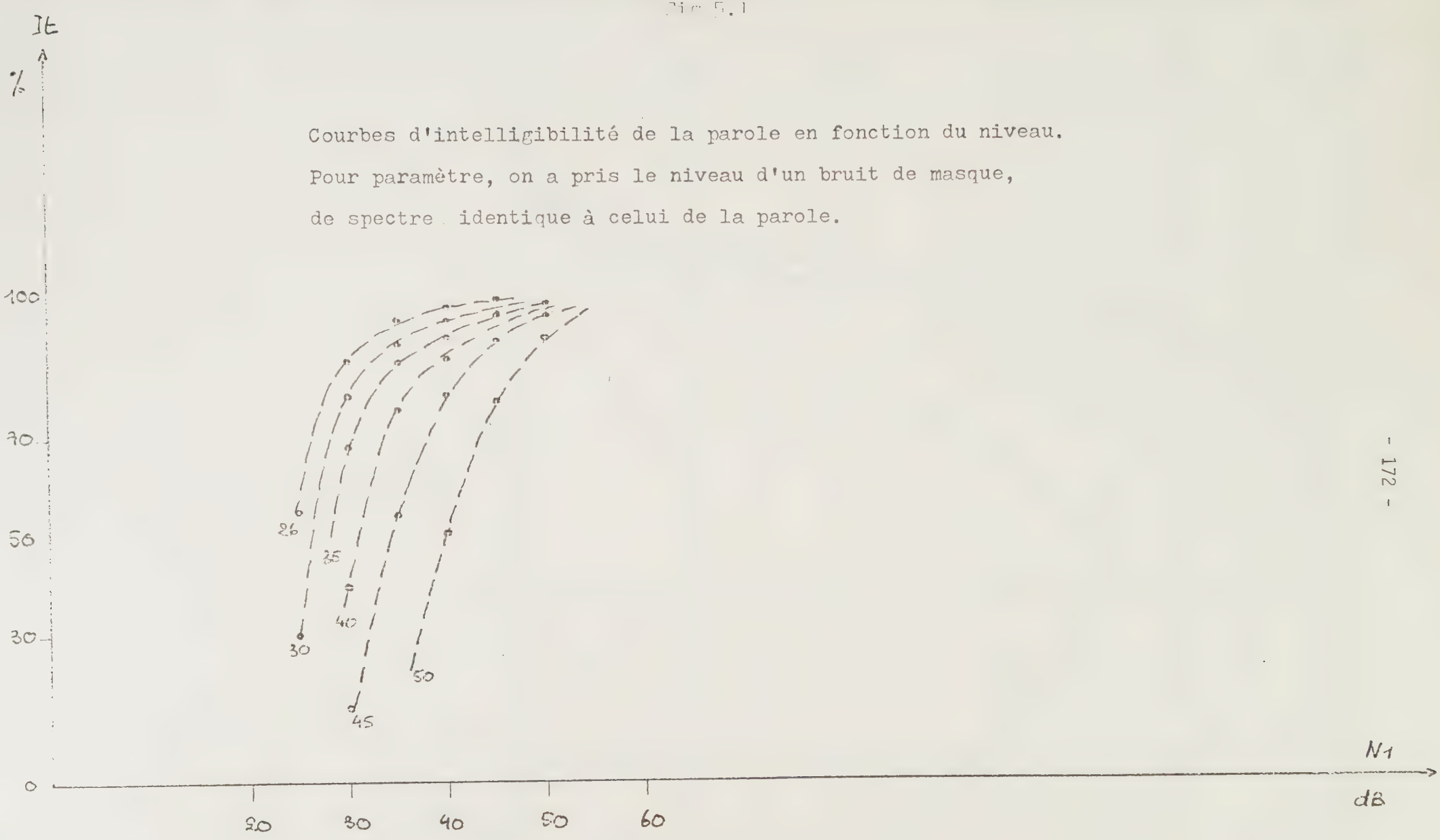
La courbe de la Fig. 5.1 a été relevée en chambre sourde dans les conditions suivantes : le signal de parole avait une direction frontale, alors que le bruit de masque était tridirectionnel : un coffret de haut-parleurs avait la même direction que le son direct, alors que les deux autres étaient à 60° à droite et 60° à gauche, dans le plan horizontal. Nous appellerons N_1 le niveau, en décibels, du signal de parole et N_B celui du bruit de masque. Nous aurons l'occasion de revenir plus loin sur l'intérêt d'avoir ainsi des courbes "moyennes", où le bruit provient simultanément de plusieurs directions : nous aurons l'occasion de voir que la direction du bruit peut provoquer des différences d'intelligibilité très sensibles, et que la situation que nous avons envisagée est située à mi-chemin entre la plus favorable et la plus défavorable.

De l'ensemble des courbes de la Fig. 5.1, on peut tirer les conclusions suivantes : l'intelligibilité croît rapidement avec N1, jusqu'à un certain niveau, où les variations sont beaucoup plus lentes. Si, par conséquent, on veut mettre en lumière des différences sensibles d'intelligibilité, d'une place à l'autre d'une même salle, on aura tout intérêt à utiliser des niveaux d'émission relativement faibles, en particulier si N1 - NB est supérieur à 0, comme le montrent les courbes de la Fig. 5.2 : ces courbes ont été déduites par Lochner d'un réseau de courbes donnant l'intelligibilité en fonction de N1, à NB constant, réseau analogue à celui de la Fig. 5.1.

5.2.- Les mesures subjectives d'intelligibilité.-

Nous avons effectué des mesures d'intelligibilité en utilisant

Courbes d'intelligibilité de la parole en fonction du niveau.
Pour paramètre, on a pris le niveau d'un bruit de masque,
de spectre identique à celui de la parole.



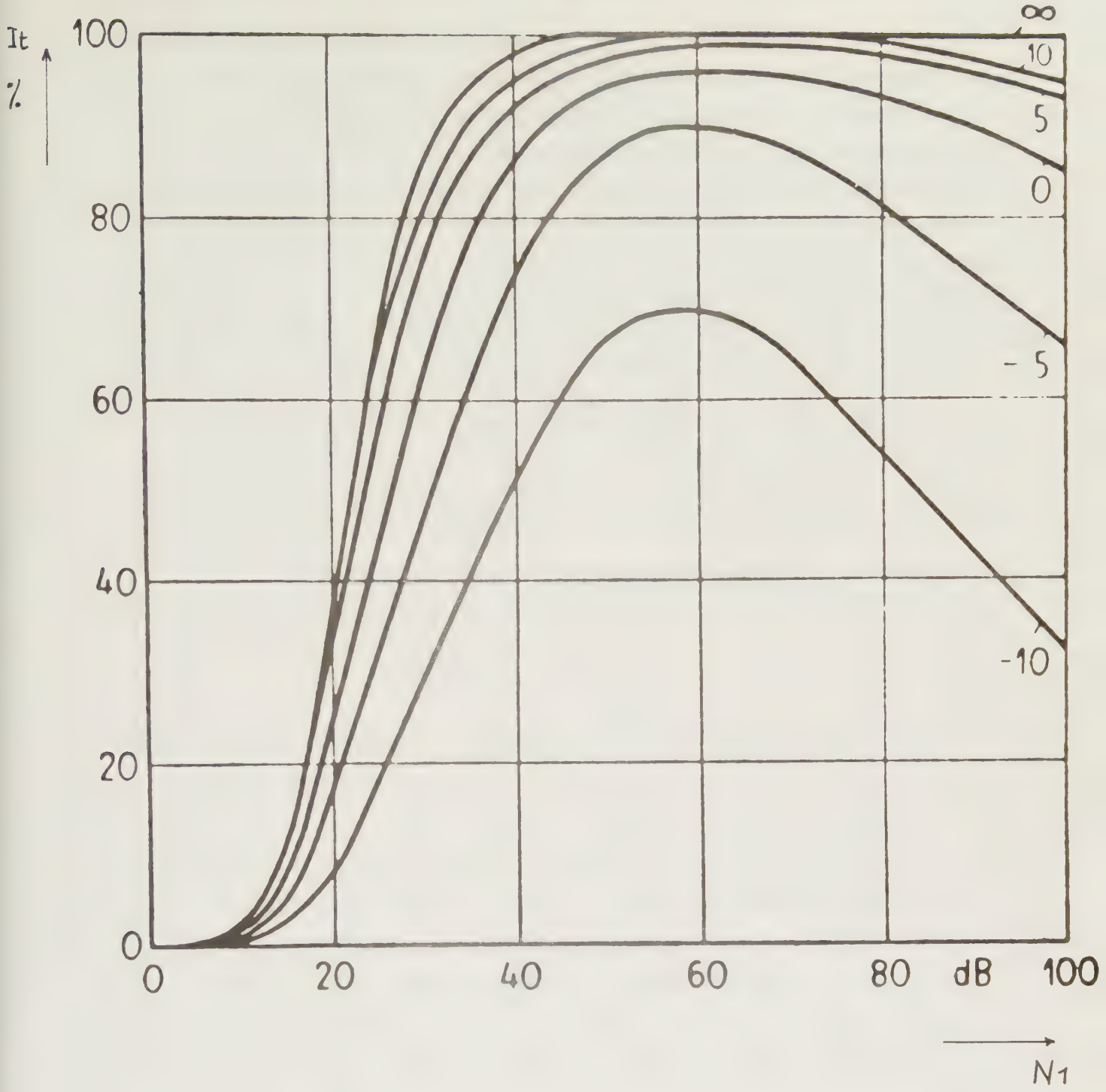


Fig. 5.2
Courbes It (N1) à N1-N2 constant,
d'après Jochner

l'ensemble des listes de logatomes dont nous avons parlé précédemment. Ces listes, préalablement enregistrées en chambre sourde, étaient diffusées à l'aide d'un émetteur omnidirectionnel constitué par un ensemble de six haut-parleurs montés sur une sphère en matière plastique, ceci pour être placé dans les mêmes conditions expérimentales que lors de l'enregistrement des échogrammes, où l'on utilisait une source également isotrope. Chaque mesure d'intelligibilité a été effectuée, pour les cinq places envisagées, numérotées de 1 à 5 (voir les Fig. 4.35 et 4.36) et portait sur l'ensemble des 1000 logatomes. Le niveau de référence était mesuré à 25 cm de la source émettrice, ce qui permettait de connaître le niveau à 1m, en champ libre.

Nous avons utilisé trois niveaux d'émission supposés mesurés à 1m : 60 dB, 50 dB et 40 dB. L'inconvénient présenté par l'utilisation de niveaux très faibles est la grande sensibilité aux bruits extérieurs, qui malheureusement ne sont pas négligeables, l'isolation de cette salle étant médiocre.

Nous présentons ci-dessous l'ensemble des résultats relatifs aux mesures d'intelligibilité effectuées en différents endroits de la salle.

	Places				
Niveaux en dB	R1	R2	R3	R4	R5
40	85%	89%	88%	88%	87%
50	90,5%	94%	92%	92%	91,5%
60	93%	94%	94%	92%	93%

On constate bien que les résultats sont d'autant plus proches que les niveaux sont plus élevés : le niveau de 40 dB semble assez faible, et laisse trop d'influence aux perturbations extérieures pour que ces résultats puissent paraître très significatifs.
Quant à ceux obtenus avec un niveau de 60 dB, ils sont très proches, et ne

mettent pas en évidence une différence significative de l'intelligibilité obtenue pour diverses places.
A titre indicatif, nous donnerons les variations de l'écart σ, en valeur absolue, de la mesure d'intelligibilité relative aux i premiers essais, à la mesure d'intelligibilité effectuée sur les cinq essais, et considérée comme mesure exacte (Fig. 5.3). Chaque essai représente une audition de 200 logatomes. Pour 4 situations de l'auditeur sur les 5 envisagées, l'écart σ décroît régulièrement lorsque le nombre d'essais augmente. Toutefois, pour la cinquième situation, l'évolution de σ est irrégulière.

5.3.- Prévisions objectives d'intelligibilité.-

a) Méthode hybride.-

Cette méthode que nous proposons est basée, d'une part sur la netteté au sens de Thiele, d'autre part sur l'utilisation des courbes de la Fig. 5.1, procédé utilisé également par Lochner.
La netteté correspond à une évaluation schématique de l'énergie utile et de l'énergie nuisible : l'énergie utile E_u correspond à l'énergie cumulée 50 ms après l'arrivée du son direct, l'énergie nuisible à l'énergie totale diminuée de l'énergie utile. Les mesures de Niese, Lochner ou Januska relèvent également d'une évaluation de l'énergie utile et de l'énergie nuisible, partielle ou totale.
Partielle pour Niese, qui tient compte de la durée de réverbération par un autre coefficient incorporé dans la formule de Knudsen; totale pour Lochner ou Januska qui estiment toute l'énergie nuisible par un procédé plus élaboré, mais semblable dans sa nature à celui de Thiele.
Nous aurons l'occasion de discuter sur la validité de l'évaluation de Thiele : Pour l'instant, admettons l'intérêt de cette distinction et son exactitude.
Appelons E_d et E_u les énergies du son direct et du son utile au niveau du récepteur : E_u sera la valeur de w(t) 50 ms après l'arrivée du son direct.
Si D est la distance qui sépare le récepteur de la source, le niveau du son utile sera :

$$N_u = N_o - 20 \log D + 10 \log (E_u/E_d)$$

où N_o est le niveau du son direct à 1m, fixé à 40, 50 ou 60 dB. Le rapport

nombre d'essais.
1
2
4
5

E_u/E_d est évalué à partir de l'échogramme calculé ou enregistré. Du niveau N_u , on déduit, à partir du rapport E_n/E_u, également connu d'après la courbe w(t), le niveau de l'énergie utile N_n. On en tire alors à l'aide du réseau de courbes de la Fig. 5.1 une estimation de l'intelligibilité.

Nous résumerons par le tableau ci-dessous, pour les cinq situations envisagées, les résultats de l'estimation d'intelligibilité It_c, lorsque le son direct à 1m a un niveau N_o de 50 dB. N_d est le niveau du son direct pris au récepteur.

	R1	R2	R3	R4	R5
N_d (dB)	34,5	29	36	29	37
N_u	42	43,5	42,5	43,5	44,5
N_n	41	37,5	40,5	38,5	38,5
It_c %	89	93	90	92	93
It_e %	90,5	94	89,5	92	92,5

Les résultats obtenus sont très satisfaisants, on peut même avancer qu'il semblerait difficile d'obtenir mieux. Nous allons maintenant examiner les résultats donnés par d'autres évaluations des énergies utile et nuisible, ou encore par l'utilisation d'une formule.

b) Evaluation de Thiele-Januska [3.11]

En reprenant l'évaluation de la netteté de Thiele, Januska a proposé une estimation d'après la formule suivante :

$$It = k_c \cdot k_1$$

où k_1 est un coefficient (inférieur ou égal à 1) lié au niveau intrinsèque de l'énergie utile, alors que k_c est déduit de la netteté C, où

$$C = E_u / (E_u + E_n)$$

par la formule empirique : $k_c = \sqrt[8]{C}$. A supposer que l'on prenne $k_1 = 1$, on obtient les résultats suivants :

Situation du récepteur	R1	R2	R3	R4	R5
C	0,55	0,8	0,61	0,75	0,73
It %	86	94	88	93	92

Ec comparant ces résultats à ceux de la méthode hybride proposée ci-dessus, on peut voir que l'estimation de Thiele-Januska varie bien dans le même sens. Elle semble cependant légérement pessimiste et d'autant plus que le coefficient k_1 correspondant à N_o = 50 dB est inférieur à 1.

c) Evaluation de Niese [3.33]

Nous avons utilisé la méthode indiquée par Niese, en calculant un degré d'écho. Nous avons au préalable divisé l'échogramme en tranches de 10 ms. L'application numérique a été effectuée pour la situation R1.

L'énergie utile est :

$$E_u = \int_{t=0}^{30ms} |p(t)|^2 \, dt = 38,2 \quad \text{(les unités sont arbitraires)}$$

En admettant que la durée de réverbération de la salle est de 0,7 s, la fonction $p_s(t)$ définie au paragraphe 3.2.5 c est alors :

$$p_s(t) = p_o \cdot f(t) \quad , \text{avec} \quad p_o = \sqrt{2E_u / 30} = 1,6$$

$$\text{et} \quad f(t) = \sqrt{\frac{\tau_N e^{-(t-30)/\tau_N} - \tau_o e^{-((t-30/\tau_o}}{\tau_N - \tau_o}}$$

avec $\tau_N = 51$ ms et $\tau_o = 23$ ms.

Si E est l'énergie de l'échogramme à l'instant t, l'énergie nuisible sera :

$$E_n = \int_{30ms}^{\infty} (\sqrt{E} - p_s)^2 \, dt$$

Nous avons divisé l'échogramme en tranches de 10 ms : le temps est compté à partir de l'arrivée du son direct. La quantité $\sqrt{E} - p_s$ est considérée comme nulle si $p_s > \sqrt{E}$. D'autre part, on prend $S = (\sqrt{E} - p_s)^2$. On obtient alors le tableau porté en annexe 1, et l'on en déduit une valeur du degré d'écho :

$$\varepsilon = E_n / (E_n + E_u) \quad , \text{avec} \quad E_n = 12,44 \quad \text{et} \quad E_u = 38,2$$

on obtient alors $\varepsilon = 0,24$ et par conséquent $k_E = \sqrt[4]{0,76} = 0,93$

Le résultat obtenu semble satisfaisant, puisque l'intelligibilité est le produit de k_E par d'autres facteurs, en particulier le facteur de niveau, et celui associé à la durée de réverbération, tous ces facteurs étant inférieurs à 1.

d) Evaluation selon la méthode de Lochner et Burger [3.30]

Comme pour l'évaluation précédente, nous avons divisé l'échogramme en tranches de 10 ms, pour la même situation R1.
Nous avons regroupé dans l'annexe 2 les détails de ce calcul : l'énergie de chaque tranche et sa contribution à l'énergie utile, et l'évaluation de l'énergie nuisible, avec des unités arbitraires. Nous arrivons à un rapport entre énergie utile et énergie nuisible de 5,2 dB. Ce rapport est nettement plus élevé que celui de la netteté de Thiele, utilisé plus haut, et qui mettait en évidence une différence de 1 dB seulement : on en déduit une prévision d'intelligibilité de l'ordre de 95%, ce qui semble nettement élevé par rapport au résultat que nous avions proposé avec la méthode hybride. Notons que la différence entre cette estimation et les autres est due à une estimation de l'énergie nuisible nettement inférieure à celle préconisée par les autres méthodes.

e) Conclusion des prévisions d'intelligibilité.-

Nous avons procédé ci-dessus à quelques estimations d'intelligibilité suivant diverses méthodes, qui permettent de se ramener au problème classique de l'intelligibilité en présence de bruit. Nous n'envisagerons pas certaines méthodes basées sur les échogrammes, mais beaucoup plus sophistiquées, et dont les chances de réussite nous paraissent à priori plus faibles, et en particulier celle du "centre de gravité" de Kürer.

Parmi les diverses méthodes envisagées, nous donnerons une préférence à notre méthode hybride qui semble, dans l'exemple choisi, suivre d'assez près les variations de l'intelligibilité mesurée : celle de Niese semble plausible, et l'on peut considérer que l'évaluation de Niese-Januska doit elle aussi donner des résultats satisfaisants. Par contre, nous émettrons quelques réserves sur l'estimation de Lochner dont le bilan donne une énergie utile relativement élevée par rapport à l'énergie nuisible et par conséquent une prévision d'intelligibilité trop optimiste.
Etant donné que notre mesure d'intelligibilité semblait déjà élevée, à cause du nombre relativement réduit de logatomes, et de l'utilisation d'un auditeur entraîné, il apparaît que l'estimation de Lochner s'écarte d'autant plus sensiblement des autres, ainsi que des mesures subjectives.

Ceci dit, nous avons procédé à ces estimations à titre d'exemple. En effet, nous pensons qu'un échogramme uniquement temporel ne rend compte qu'en partie seulement de la réalité. On sait que l'audition est un phéno-

mène spatial, étant donné la présence des deux oreilles. Les effets de l'audition binaurale sont maintenant entrés dans les moeurs, avec l'utilisation de la stéréophonie, puis de l'ambiophonie. Si l'on examine l'allure des échogrammes relatifs à la salle de conférences, donnés par les Fig. 4.43 et 4.44, on remarque que le son direct est suivi de réflexions relativement proches, et qui, entre elles, sont très serrées : à supposer que certaines réflexions apportent de l'énergie utile, d'autres de l'énergie nuisible, il apparaît bien que les unes et les autres proviennent à peu près de toutes les directions.
Etant donné la proximité des différentes réflexions, on peut estimer que le rôle de la direction a un effet réduit et que l'énergie utile, aussi bien que l'énergie nuisible, est diffuse.

Par contre, dans le cas d'une grande salle, les réflexions sont beaucoup plus espacées : c'est alors que l'influence de la direction peut se faire sentir. L'étude que nous allons mener dans les paragraphes suivants comporte deux volets : le premier est relatif au "seuil de perception d'un écho", c'est-à-dire au seuil au-dessus duquel une réflexion apparaîtra comme perturbatrice, donnant un "effet d'écho" par rapport au rayonnement d'une source primaire constituée par le son direct seul, ou bien par le son direct auquel on aura associé une ou plusieurs réflexions n'apportant aucun effet d'écho. Ce seuil $N_S(\tau)$, où τ est le retard de la réflexion envisagée, permettra de délimiter l'énergie utile et l'énergie nuisible de la réflexion de niveau I : si $I > I_s$, alors l'énergie $I - I_s$ sera à mettre au compte de l'énergie nuisible, l'énergie I_s correspondant à l'énergie utile de la réflexion considérée.

Le deuxième volet est une vérification de cette hypothèse : on effectue des mesures d'intelligibilité en présence d'un écho perturbateur, puis on remplace cet écho par un bruit de même spectre que la parole, dont le niveau est fixé d'après celui du seuil de perception de l'écho.

Avant d'aborder cette étude, nous donnerons quelques indications sur l'installation qui permet ce type de mesure.

5.4.- Le champ sonore synthétique.-

Comme son nom l'indique, le "champ sonore synthétique" a pour objet de reconstituer dans une chambre sourde un champ sonore, tel qu'il pourrait exister, en principe, dans une salle réelle. Plus modestement, on peut dire qu'il s'agit d'un dispositif permettant d'étudier l'effet des échos, ou d'une autre perturbation, ou encore de l'étendue de la source, sur un signal acoustique. Le travail le plus important, en ce qui concerne le rôle des échos sur la parole, reste celui de Haas, malgré son ancienneté [3.28] . Il faut noter qu'une revue américaine publiait, en 1972, une traduction de ce travail, universellement connu et cité [5.5] .
Les travaux de Haas sont relatifs à une seule réflexion. La variable utilisée est le pourcentage P d'auditeurs gênés par l'écho, Haas ayant pu utiliser le jugement de 100 auditeurs. Le "seuil de gêne" ou de "nuisance" d'un écho correspond à P = 50. Une telle méthode de mesure est appelée "méthode de la constante".

Les mesures de Haas sont relatives au pourcentage d'auditeurs gênés par un écho de retard variable. Comme paramètre, Haas a utilisé d'une part le débit de la parole soumise à la mesure, et d'autre part le niveau de la réflexion par rapport à celui du son direct. Haas a également utilisé un filtrage passe-bas ou passe-haut pour étudier l'effet du timbre sur la perception de l'écho.

En général, les mesures de Haas ont été faites en utilisant les directions suivantes : le son direct est à 45° à droite de l'auditeur, ou des auditeurs, la réflexion à 45° à gauche. Cette situation particulière a été choisie, dit Haas, car elle permet d'équilibrer facilement la sonie des deux voies. Mais Haas a également envisagé le problème de la direction de l'écho, en utilisant une direction frontale pour le son direct. Nous reviendrons sur les résultats qu'il a obtenus.

A côté du travail de Haas, on peut mentionner quelques publications qui apportent un complément d'information. En 1953, Muncey et ses collaborateurs [5.14] étendaient les manipulations de Haas au cas de signaux divers, et à plusieurs types de musique en particulier.
En 1962, Karaskiewicz et Kwiek [5.15] [5.16] reprenaient le problème

du spectre de l'écho, et plus récemment Yanagisawa et Nimura 5.17 relient la perception de l'écho à une échelle psychoacoustique de "qualité" d'une salle.

Dans le cas particulier de la parole, c'est Seraphim [5.6] [5.7] qui, en 1961, a donné un prolongement important au travail de Haas, mais dans un sens cependant différent. Le signal mesuré n'est plus le même : il ne s'agit pas du "seuil de nuisance" d'une réflexion, que l'on peut appeler également "seuil de perception de l'écho", mais du "seuil de perception absolue" d'une réflexion.

Nous reviendrons plus loin sur la définition de ce seuil. L'intérêt de cette étude est de permettre d'éliminer, dans un échogramme, les réflexions dont l'effet est négligeable, et par conséquent d'en réduire la complexité aux seules réflexions dont l'influence est sensible.

Le travail récent de Plenge [5.18] apporte une contribution dans le même sens. Les diverses mesures de Seraphim ont été reprises ou utilisées par les auteurs britanniques Marshall et Barron, pour l'étude des salles de concert [5.23] [5.24] [5.25]
D'autre part, dans plusieurs publications, Meyer et Damaske [5.8] [5.19] ont également utilisé le seuil de perception absolu, en liaison avec le degré de cohérence de plusieurs signaux.

Naturellement, le seuil de perception absolu est situé, en général, au-dessous du seuil de nuisance d'une réflexion. Il peut lui être égal, le cas échéant, par exemple si la réflexion envisagée est très retardée, si bien qu'aucun effet d'intégration au son direct n'est possible. Mais l'étude du seuil de perception absolu n'apporte pas de contribution directe au problème de l'intelligibilité de la parole.

Le travail le plus récent à notre connaissance et qui traite du problème de l'intelligibilité en présence d'échos de direction variable est du à Lehmann et Wettschureck [5.9] . Mais encore n'est-il que très partiel, et se borne-t-il à constater des variations d'intelligibilité suivant l'angle d'incidence d'un bruit de masque.

Il n'en restait pas moins qu'une étude systématique sur le seuil de perception d'un écho en présence de plusieurs réflexions restait inexistante, et qu'à notre avis, il était utile de poursuivre le travail de Haas en le généralisant au cas de réflexions multiples et de directions également multiples.

L'étude que nous allons exposer consiste à tracer des courbes de seuil de perception d'un écho, en fonction de son retard, de son intensité, et de sa direction. Mais contrairement au cas traité par Haas, le son direct ne sera pas forcément isolé, mais on lui associera des réflexions antérieures à l'écho étudié, de situations diverses. On sera fatalement amené à se limiter à un certain nombre d'exemples, mais on tâchera de les choisir judicieusement.

Nous utiliserons ces courbes pour évaluer l'énergie utile et l'énergie nuisible des réflexions qui se trouvent dans un échogramme : on pourra alors en déduire une prévision de l'intelligibilité de la parole dans les salles, que nous espérons plus satisfaisante que celles proposées à ce jour.

5.5.- Mesure du seuil de perception des échos.-

5.5.1.- Montage utilisé et mode opératoire.-

L'élément principal est constitué par un magnétophone à échos construit au laboratoire par Th. Angélini et J. Mège [5.11] [5.12] Ce magnétophone est doté de 10 têtes lectrices continûment réglables et peut donner des délais situés entre 14 et 500 ms. (Fig. 5.4) Chaque voie alimente un ou plusieurs coffrets de haut-parleurs, disposés sur une armature métallique, dans une chambre sourde (Fig. 5.5). Cette "ombrelle" matérialise, par ses arcs méridiens, une demi-sphère de 2m de rayon, centrée sur la tête d'un auditeur : 12 quarts de cercles verticaux sont répartis tous les 30° et viennent s'amarrer sur un grand cercle horizontal. Chaque coffret, équipé de 3 haut-parleurs disposés en trèfle, est en outre muni d'un pavillon exponentiel, ce qui lui assure une bonne directivité vers le centre de la sphère. L'azimut θ peut prendre des valeurs variables de 30° en 30°, le site φ est continûment variable.

Le signal de parole qui sert à la mesure est le mot "acoustique" enregistré sur une boucle sans fin, avec un débit de 240 ms par syllabe : un son à 1000 Hz, dont le niveau est à 6 dB au-dessous du maximum du signal de parole, permet d'évaluer le niveau utilisé.

Le mode opératoire est le suivant : la "source primaire" est constituée par le son direct seul, ou le son direct et une ou plusieurs réflexions qui lui sont parfaitement intégrées, c'est-à-dire ne provoquent aucun effet d'écho. L'intensité de la source primaire est égale à la somme de l'intensité du son direct et des réflexions associées. A cette source primaire, on juxtapose un écho dont on va étudier le seuil.

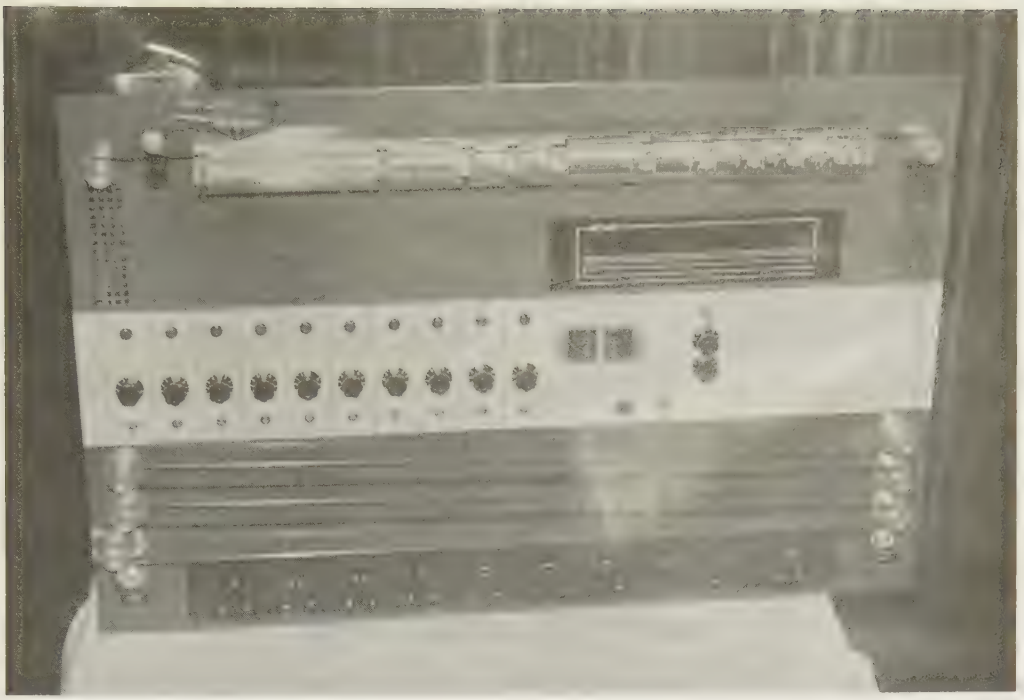


Fig. 5.4 - Le magnétophone à échos -

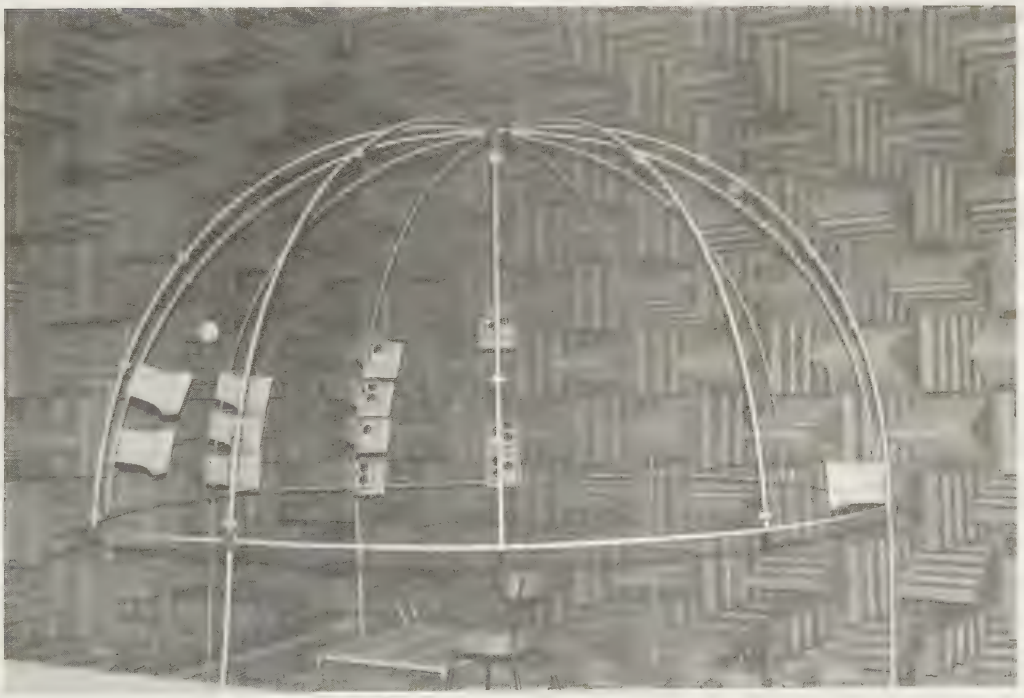


Fig. 5.5 - L'ombrelle et les coffrets de haut-parleurs -

L'intensité et le niveau de la voie i sont notés I_i et N_i. Le son direct est donné par la voie 1, son niveau sera N_1, son intensité I_1. Les directions angulaires θ_i et φ_i de la voie i sont repérées conformément aux schémas des figures 5.12 et 5.14.

a) - Définition du seuil de perception de l'écho.-

Nous n'avons pas utilisé comme Haas une méthode de mesure par la "moyenne", mais une "méthode d'ajustement". Etant donné la source primaire avec son intensité et sa direction, et une réflexion dont le retard et la direction sont également fixés, on fait décroître l'intensité de cette réflexion à partir d'une valeur assez haute pour qu'un effet d'écho soit sensible : cet effet disparaît pour une valeur n du niveau N de la réflexion.
De même, en partant d'une valeur très faible de N, on l'augmente jusqu'à ce qu'un effet d'écho devienne perceptible, ce qui a lieu pour une valeur n' de N. Le seuil de perception de l'écho est la valeur $N_S = (n + n')/2$.
La précision de ces mesures peut être estimée à $\pm$ 2 dB : elle dépend d'une part de l'appréciation de l'auditeur, de la constance de ses jugements, et d'autre part, de la précision des potentiomètres qui réglent le niveau des différentes voies. Les graduations des boutons des potentiomètres correspondent à des intervalles de 4 dB. Le niveau des différentes voies est mesuré à l'aide d'un sonomètre dont le micro est situé au centre de la coupole, c'est-à-dire à la place de la tête de l'auditeur. Il faut noter qu'on demande à l'auditeur de faire abstraction des effets d'élargissement de la source sonore ou d'impression spatiale qui font d'ailleurs l'objet d'autres travaux relatifs à des salles de concert [5.20] [5.21] [5.22] .

b) - Définition du délai critique de Haas.-

Etant donné le son direct et une réflexion d'intensités et de directions données, le délai critique τ_c employé par Haas correspond au délai de la réflexion pour lequel 50% des auditeurs interrogés s'estiment gênés par un effet d'écho.

c) - Comparaison des deux méthodes.-

Deux questions se posent alors : pourquoi avoir choisi cette méthode d'ajustement plutôt que la méthode de la moyenne employée par Haas? d'autre part, le seuil mesuré donne-t-il les mêmes résultats que le "délai critique" défini par Haas, correspondant à P = 50 ?

Nous avons utilisé la méthode d'ajustement, car elle est beaucoup plus économique, et temps et en personnel, que la méthode de la moyenne. Haas a utilisé cent auditeurs qui ont donné leur avis sur la nuisance de l'écho. On conçoit facilement la longueur d'une mesure nécessitant la réaction de cent personnes, si l'on envisage de les soumettre une à une à l'essai. Il nous était difficile de rassembler plusieurs auditeurs sur la plateforme exiguë que nous utilisions dans la chambre sourde; d'autre part, le fait d'utiliser plusieurs auditeurs situés au centre d'une coupole de dimensions restreintes, c'est-à-dire de deux mètres se rayon seulement, risquait d'entraîner des différences sensibles entre les situations acoustiques des divers auditeurs. D'ailleurs, les mesures que nous voulions effectuer, c'est-à-dire les variations du seuil de perception d'un écho avec son délai auraient demandé un travail expérimental gigantesque, étant donné qu'il aurait fallu tracer une courbe $P(\tau_2)$ avec I_1 et I_2 donnés pour obtenir un seul point utile, c'est-à-dire la valeur de τ_2 pour P = 50%.

La méthode que nous avons utilisée ne nécessite qu'un seul auditeur : on peut trouver qu'il y a là un inconvénient sérieux, et qu'il est hasardeux de se fier au jugement d'un seul individu. Mais nous avons pris soin de faire vérifier certains résultats par des auditeurs de bonne volonté. A ce sujet, comme pour les mesures d'intelligibilité, se pose également le problème de l'auditeur averti, car un auditeur "naïf" aura facilement tendance à confondre le seuil de perception de l'écho avec le seuil absolu de perception, en particulier si les directions du son direct et de la première réflexion sont différentes.

5.5.2. - Seuil de perception de l'écho donné par la voie 2.-

Il s'agit de la manipulation la plus simple. La source primaire est constituée par le son direct seulement, auquel on juxtapose une réflexion de retard, de direction et d'intensité variables. C'est le seul problème envisagé par Haas.
En ce qui concerne le "temps critique" envisagé par cet auteur, les mesures obtenues par la méthode d'ajustement semblent donner des résultats sensiblement plus faibles que ceux donnés par la méthode de la moyenne. Par exemple, pour les directions θ du son direct et de l'écho de 45° et -45°, le temps critique de Haas est de l'ordre de 60 ms, alors que celui que nous avons mesuré est de 40 ms seulement, ces délais étant obtenus, semble-t-il, pour une même vitesse d'élocution.

Ces différences semblent dues au fait que les auditeurs utilisés par Haas n'étaient pas spécialement entraînés, et le seuil de gêne qu'ils relevaient se trouve de ce fait au-dessus du seuil de perception que nous avons mesuré. Ces délais sont relatifs à un écho de même intensité que le son direct.

a) - Remarque sur la direction du son direct.-

Parmi les auteurs qui ont travaillé sur le problème de la perception du premier écho, il peut paraître assez curieux que Haas, aussi bien que Lochner ait envisagé à priori un son direct à 45° d'un côté, alors que la première réflexion provenait d'une direction symétrique. Nous avons mentionné la raison pratique qui avait amené Haas à choisir cette disposition : Lochner n'en donne aucune. Il semble logique de regarder la source, dont la direction est celle du son direct, c'est pourquoi nous prendrons en général $\theta_1 = \varphi_1 = 0$.

A titre de comparaison, nous avons tracé les courbes $N_{2s}(\tau_2)$ pour $\theta_1 = 45°$ et $\theta_2 = -45°$ (Fig.5.6). Le niveau du son direct était fixé arbitrairement à 71 dB.

On peut comparer cette courbe avec celles données par Lochner (Fig.5.7). La courbe que nous avons trouvée est située grosso-modo entre les deux courbes de Lochner, cependant la pente est sensiblement plus forte.

Il paraît probable que les différences observées entre les courbes de Lochner et la nôtre soient dues à la diversité des débits de la parole utilisée.

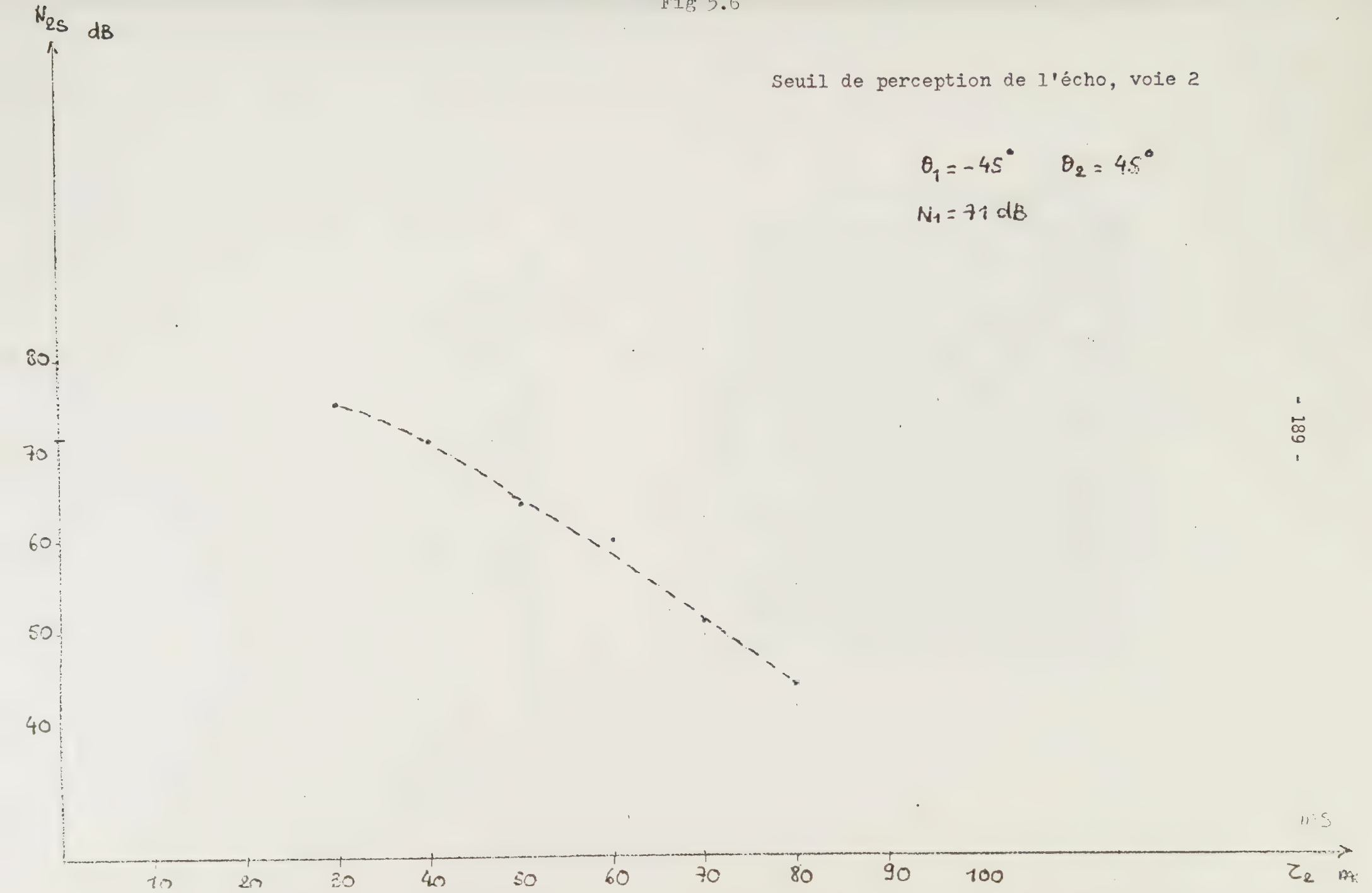
b) - Remarque sur l'intensité du son réfléchi.-

Considérons avec Knudsen [3.14] une salle où la source est située en S, le récepteur en S' (Fig. 5.8).

En général, le son correspondant à une réflexion a un niveau plus faible que le son direct, en particulier si la réflexion est provoquée par une paroi plane. L'atténuation est due, si l'on néglige l'absorption du son par l'air, au seul coefficient de réflexion de la paroi envisagée. Cependant, il peut arriver qu'un effet de focalisation donne un son réfléchi plus intense que le son direct: mais en général, cette réflexion ne provient pas d'une seule direction, mais de plusieurs, et il sera exceptionnel qu'une réflexion plus intense que le son direct provienne d'une seule direction.

Nous avons cependant envisagé ce cas particulier qui correspond, non pas au masquage d'une réflexion par le son direct, mais à l'inverse.

Fig 5.6


N_{2S} dB
Seuil de perception de l'écho, voie 2
$\theta_1 = -45°$ $\theta_2 = 45°$
$N_1 = 71$ dB
80
70
60
50
40
10
20
30
40
50
60
70
80
90
100
τ_2 ms

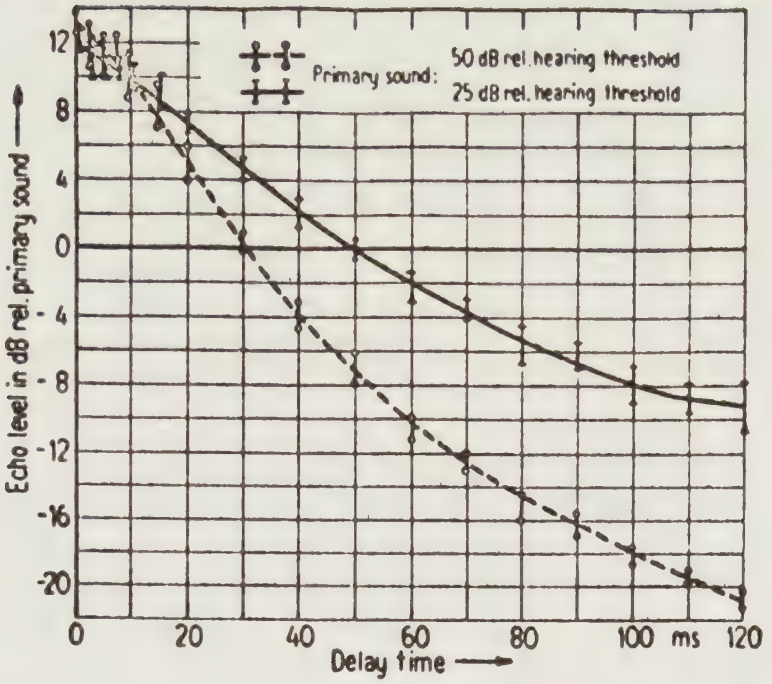
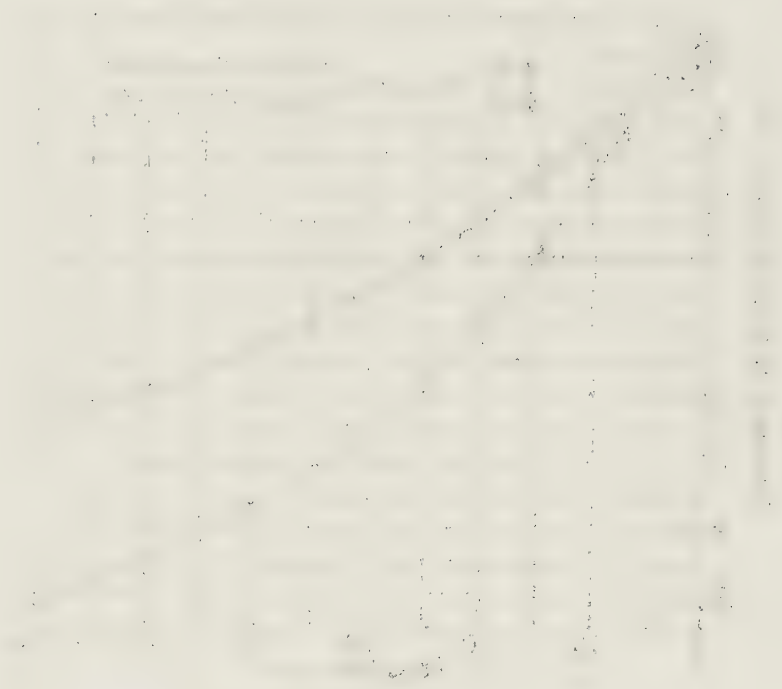


Fig 5.7 : seuil de perception de l'écho,
d'après Lochner et Burger.
directions $\theta_1 = 45°$, $\theta_2 = -45°$



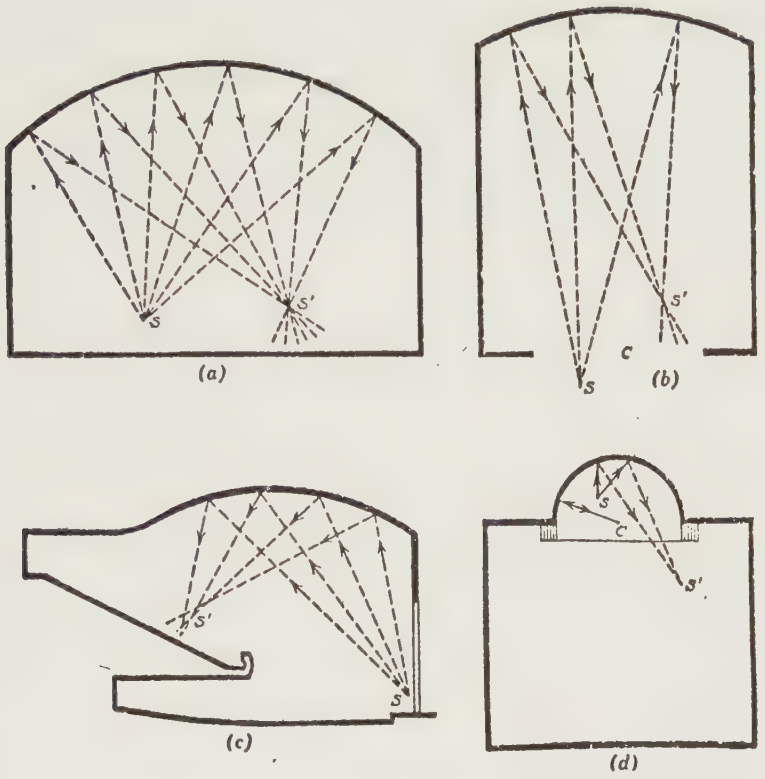


Fig 5.8. Exemples de focalisation, d'après Knudsen.

La figure 5.9 montre les résultats obtenus dans le cas où $\theta_1 = \theta_2 = 0$ et $\varphi_1 = \varphi_2 = 0$. Le niveau du son direct était de 60 dB. Le seuil S_2 mesuré correspond au seuil de disparition de l'écho: pour $N_2 > S_2$, l'effet d'écho disparaît alors qu'il est présent entre les courbes S_2 (τ_2) et N_{2S} (τ_2), la courbe N_{2S} (τ_2) étant la courbe de seuil de masquage de l'écho par le son direct. Le montage dont nous disposons ne nous permettant pas d'obtenir des délais inférieurs à 15 ms, il est apparu qu'en deçà de 30 ms, le seuil était stationnaire,aux alentours de 10 dB au-dessus du son direct: par contre, au-delà de 30 ms, le seuil augmente rapidement et atteint des valeurs très élevées.

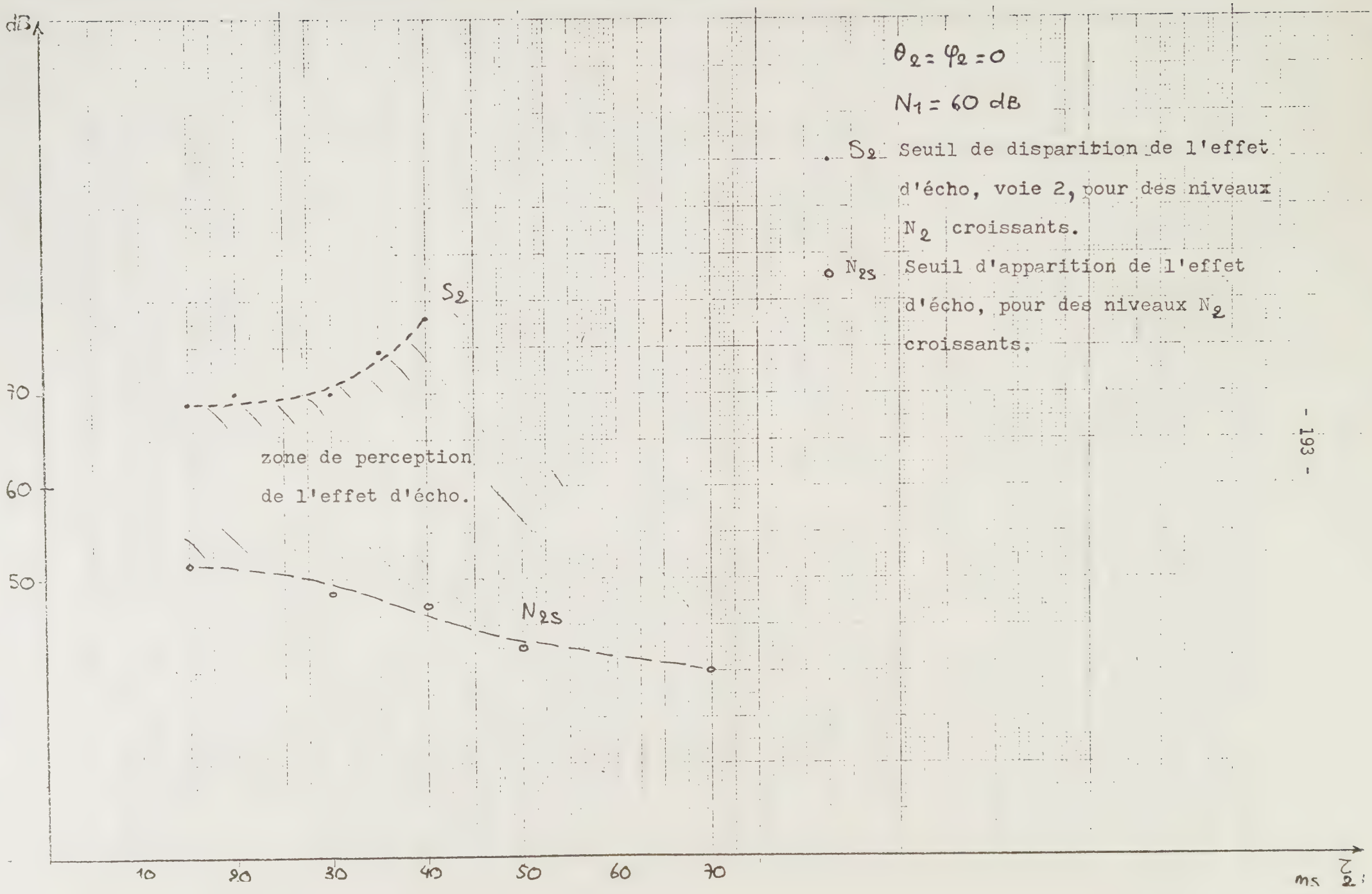
Si la direction θ_2 est différente de θ_1, ne serait-ce que de 30°, le seuil S_2 a une valeur très élevée, supérieure de 20 dB au moins à N_1, même pour des délais très faibles (τ_2 = 15 ms). On voit que le processus d'intégration auditif est éminemment chronologique, et que l'effet de "masquage inverse" du son direct par une réflexion est possible seulement si le niveau de la réflexion est considérablement plus élevé que celui du son direct, dès que les directions sont différentes. Cet effet met bien en lumière le rôle perturbateur d'une focalisation, même si la réflexion ainsi créée a un délai assez faible. Il est à noter que si un tel phénomène est assez rare dans le cas d'un processus naturel de réverbération, il peut survenir facilement avec un système de sonorisation mal réglé: si l'amplification est trop faible, le niveau du signal produit par un haut-parleur proche peut être inférieur à celui qui provient de la source réelle et amener un effet d'écho.

Les différences observées entre le cas où la réflexion a la même direction que le son direct, ou une autre, peuvent s'expliquer en disant qu'un son masque d'autant plus facilement un autre que leurs directions sont voisines. Cet effet, intuitivement, se conçoit assez bien: cependant, nous aurons l'occasion de voir par la suite qu'une telle hypothèse peut être mise en défaut. Il semble, dans le cas de l'effet de masquage inverse qui fait l'objet de ce paragraphe, que le seuil mesuré soit un "seuil de perception absolu" du son direct. Dès que le son direct est perceptible, son effet est perturbateur, ce qui explique le niveau très élevé de S_2 pour $\theta_2 = \theta_1$.

c) - Cas d'une source primaire monodirectionnelle.-

On a fixé comme direction du son direct $\theta_1 = 0, \varphi_1 = 0$ et en général, nous avons utilisé cette direction dans le cas de toutes nos mesures.

Fig 5.9


N_2
dB
$\theta_2 = \varphi_2 = 0$
$N_1 = 60$ dB
S_2 Seuil de disparition de l'effet d'écho, voie 2, pour des niveaux N_2 croissants.
N_{2S} Seuil d'apparition de l'effet d'écho, pour des niveaux N_2 croissants.
S_2
N_{2S}
zone de perception de l'effet d'écho.
70
60
50
10
20
30
40
50
60
70
ms
τ_2

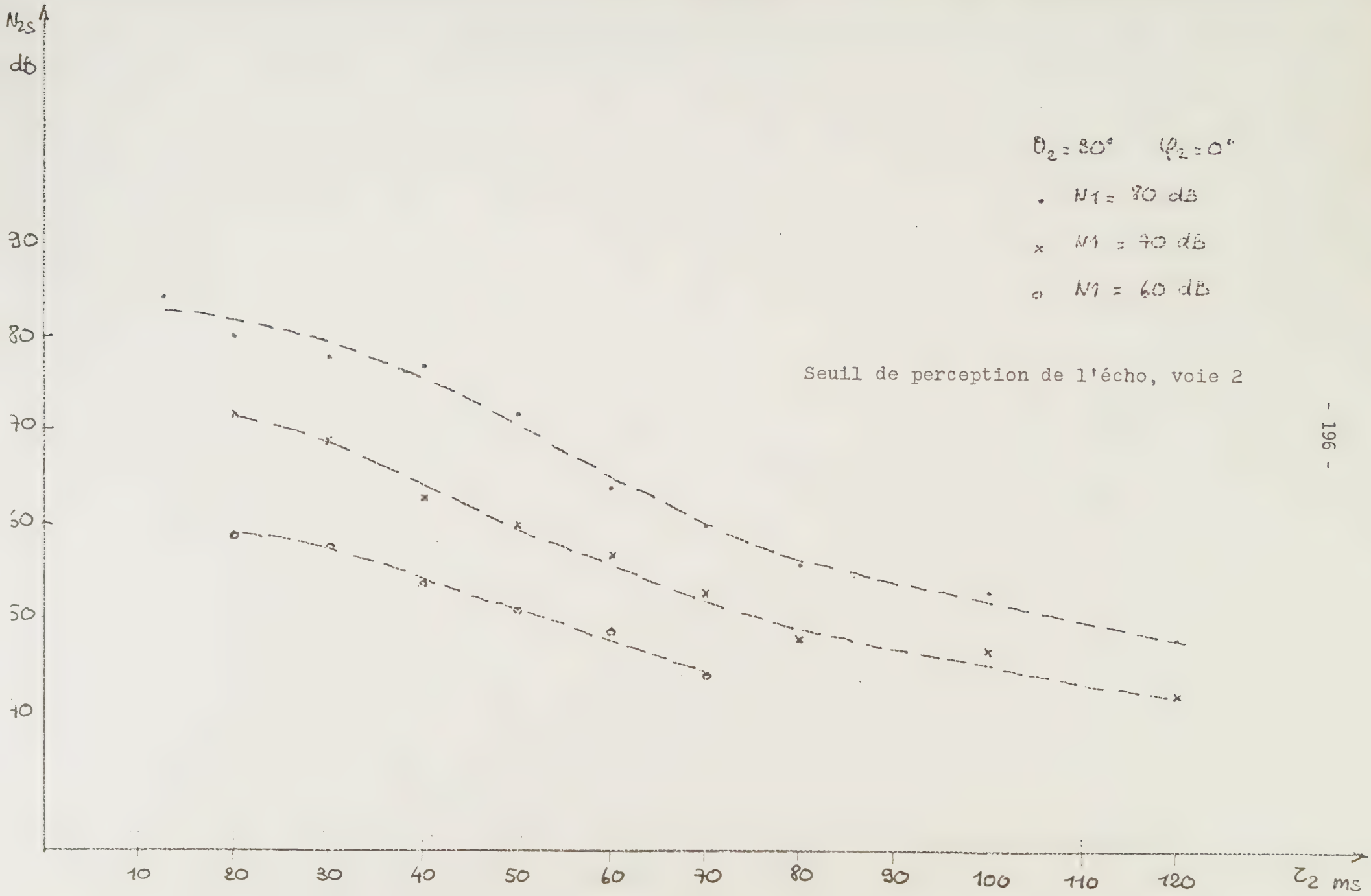
Pour $\theta_2 = 0$, $\varphi_2 = 0$, nous avons traçé les courbes $N_{2S}(\tau_2)$ pour différentes valeurs de N_1 (Fig. 5.10). Ces courbes appellent les commentaires suivants : tout d'abord, le niveau du seuil relevé, même pour des délais relativement faibles, est inférieur à celui du son direct. Ce résultat peut paraître surprenant, si l'on considère les courbes des figures 5.6 et 5.7. Il met bien en évidence que le seuil de perception dépend de la direction des sons considérés. Les courbes présentent une pente maximale entre 40 et 60 ms, de l'ordre de 0,5 dB par milliseconde. Pour différents niveaux du son direct, les courbes présentent une allure très voisine. La pente devient plus faible si N_1 décroît, cependant ses variations demeurent faibles, et moindres que celles indiquées par Lochner.

Si l'on envisage la direction $\theta_2 = 30°$, on constate que le seuil est notablement plus élevé, dans les mêmes conditions, que pour $\theta_2 = 0$. Bien entendu, la source sonore n'est plus localisée en un point, mais élargie: cependant, on demande à l'auditeur de ne pas tenir compte de cet effet, et de concentrer son attention sur la nuisance éventuelle de la réflexion envisagée. La figure 5.11 montre trois courbes enregistrées dans ces conditions, dont l'allure est très voisine de celles de la figure 5.10.
Pour les situations $\theta_2 = 60°$, 90°, 120°, et 150°, on a également tracé des réseaux de courbes qui sont en tout point très proches des résultats obtenus pour $\theta_2 = 30°$. Par contre, pour $\theta_2 = 180°$, on trouve des courbes situées de façon analogue à celles obtenues pour $\theta_2 = 0°$, et qui mettent en évidence une diminution du seuil de perception pour cette valeur particulière de l'azimut. La courbe de la figure 5.12 résume les résultats obtenus, en indiquant les variations du seuil de perception de l'écho à 30 ms. La courbe passe par un minimum pour $\theta_2 = 0°$ et $\theta_2 = 180°$, et par un palier entre ces deux valeurs.

Si l'on fait varier φ_2, en gardant un angle θ_2 constant, on constate que le seuil de perception subit peu de variations par rapport à la situation $\varphi_2 = 0°$, à θ_2 constant (Fig. 5.13) : cependant, aux alentours de $\varphi_2 = 150°$, le seuil passe par un maximum. La courbe $N_{2S}(\varphi_2)$ pour $\theta_2 = 0°$ est donnée par la Fig. 5.14. On a pris $\varphi_2 = 30$ ms.

A titre de comparaison, nous avons porté sur la figure 5.15 les résultats obtenus par Haas : la variable utilisée n'est pas le seuil de perception à 30 ms, mais le délai critique τ_o d'une réflexion de même niveau que le son direct. Nous avons déjà remarqué que le délai critique de Haas est nettement supérieur à celui déduit des courbes que nous avons tracées, mais les variations $\tau_c(\theta_2)$ ou $\tau_c(\varphi_2)$ sont analogues à celles de $N_{2S}(\theta_2)$

Fig 5.11

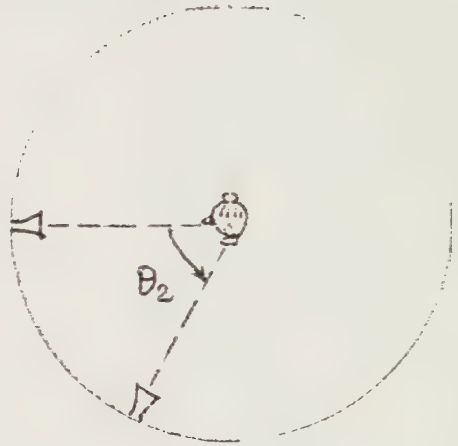

N_{2S}
dB
$\theta_2 = 30°$ $\varphi_2 = 0°$
• N_1 = 80 dB
× N_1 = 70 dB
○ N_1 = 60 dB
Seuil de perception de l'écho, voie 2
80
70
60
50
40
10
20
30
40
50
60
70
80
90
100
110
120
τ_2 ms

Seuil de perception de l'écho, voie 2.

$\varphi_2 = 0$

$\tau_2 = 30$ ms

$N1 = 70$ dB



N_{2S} dB

70

60

50

30 60 90 120 150 180 θ_2°

Seuil de perception de l'écho, voie 2.

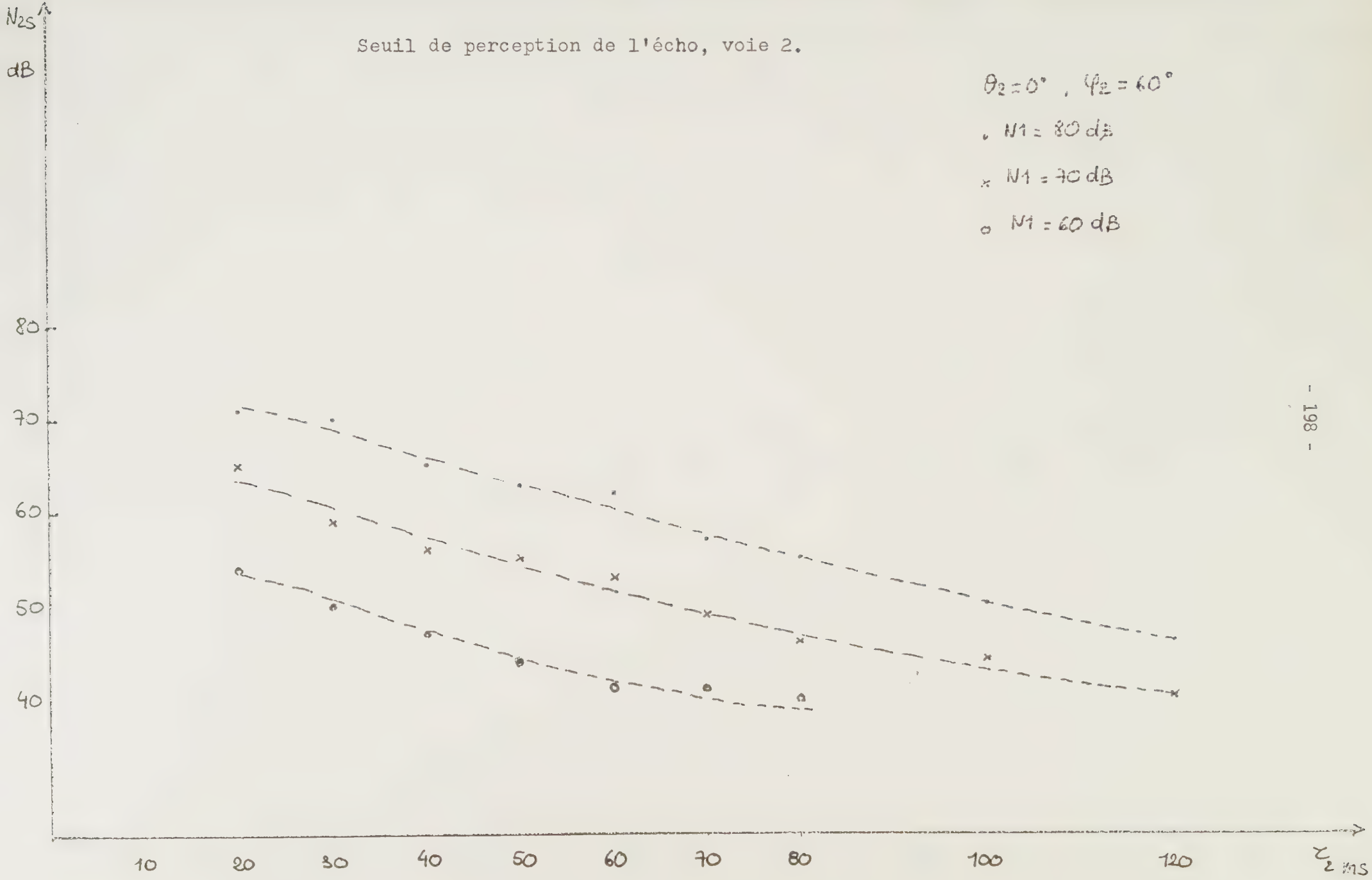


Fig 5.14

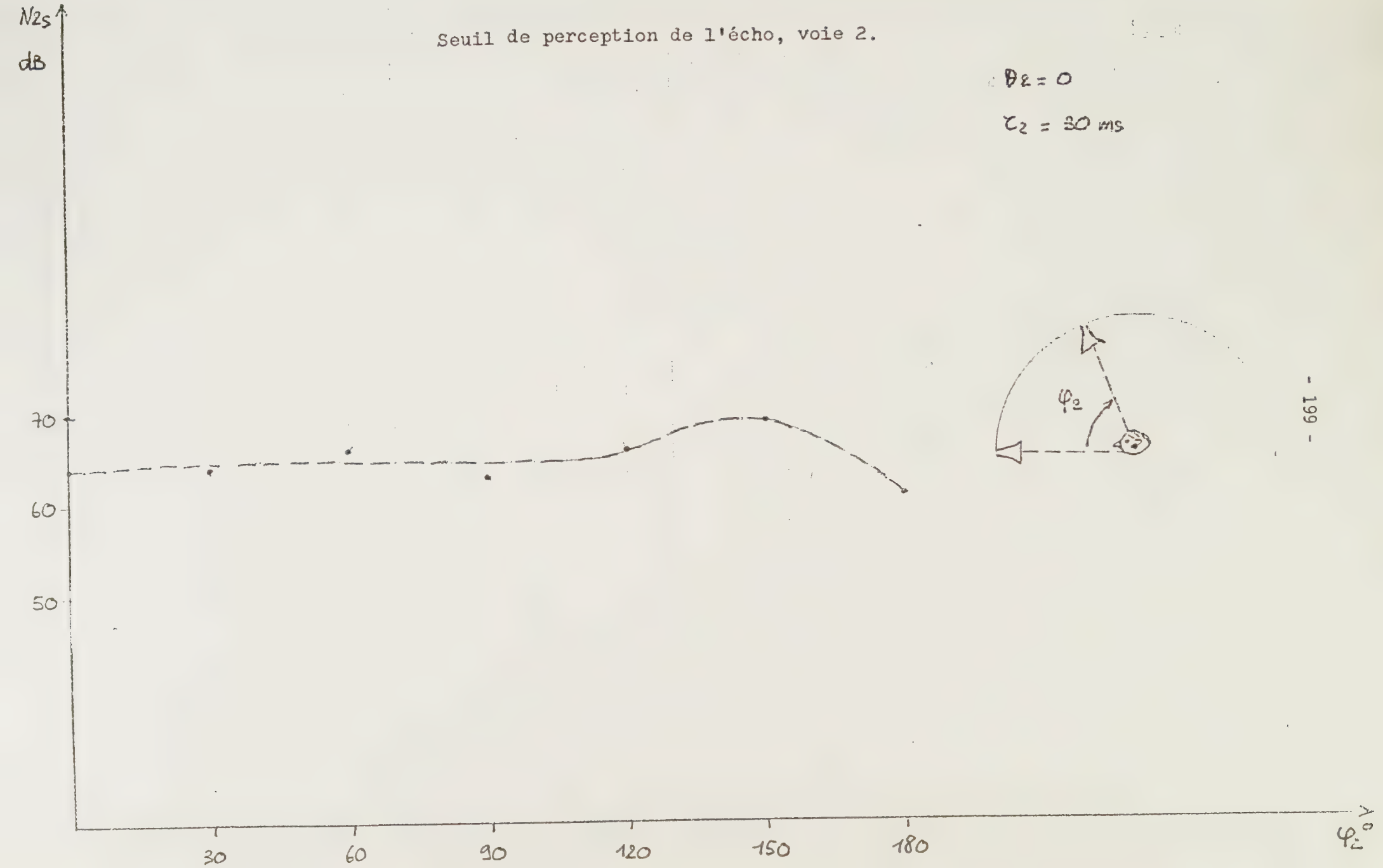
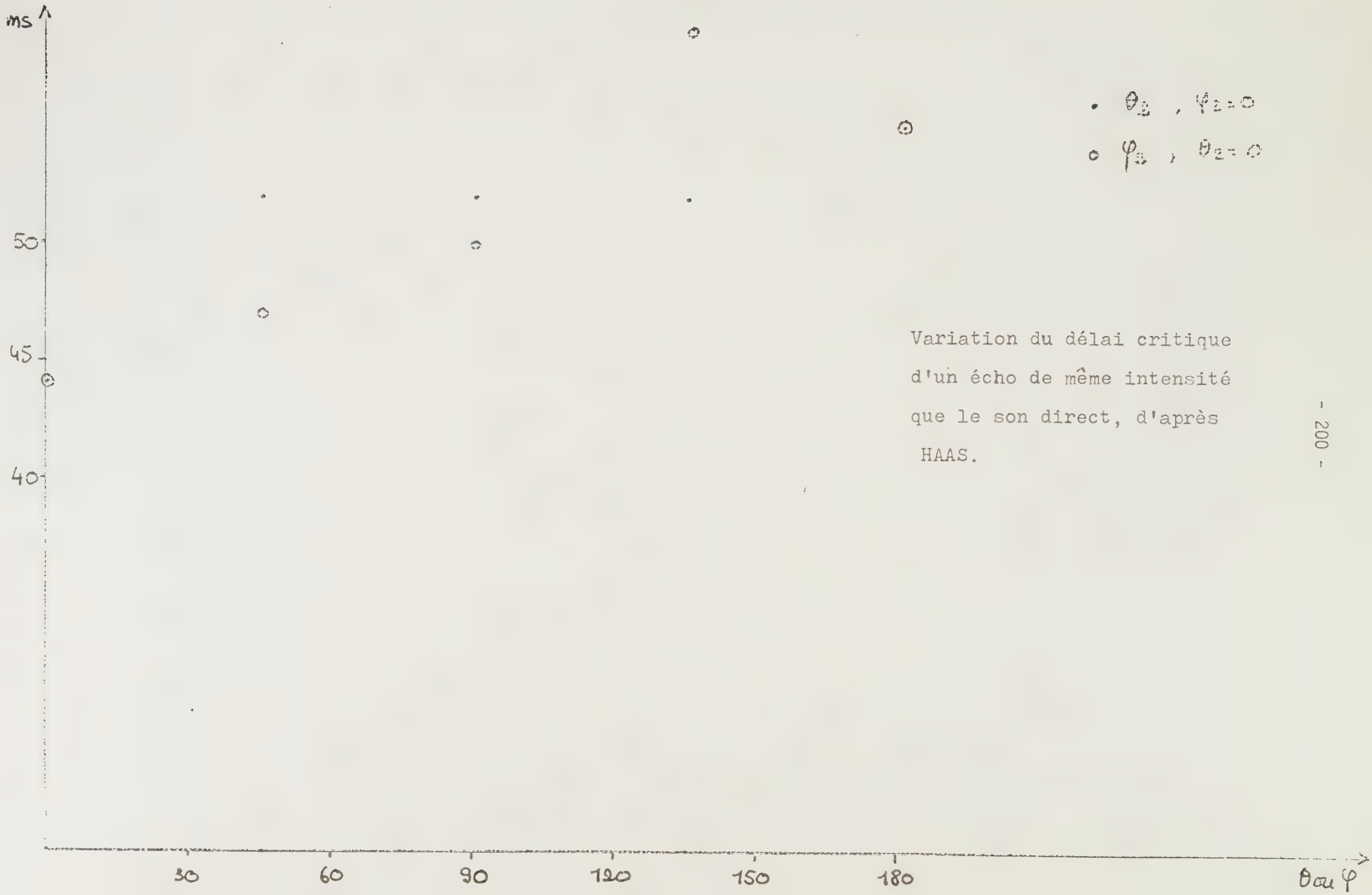

N2s
dB
Seuil de perception de l'écho, voie 2.
$\vartheta_2 = 0$
$\tau_2 = 30$ ms
φ_2
70
60
50
30
60
90
120
150
180
φ_1

Fig 5.15



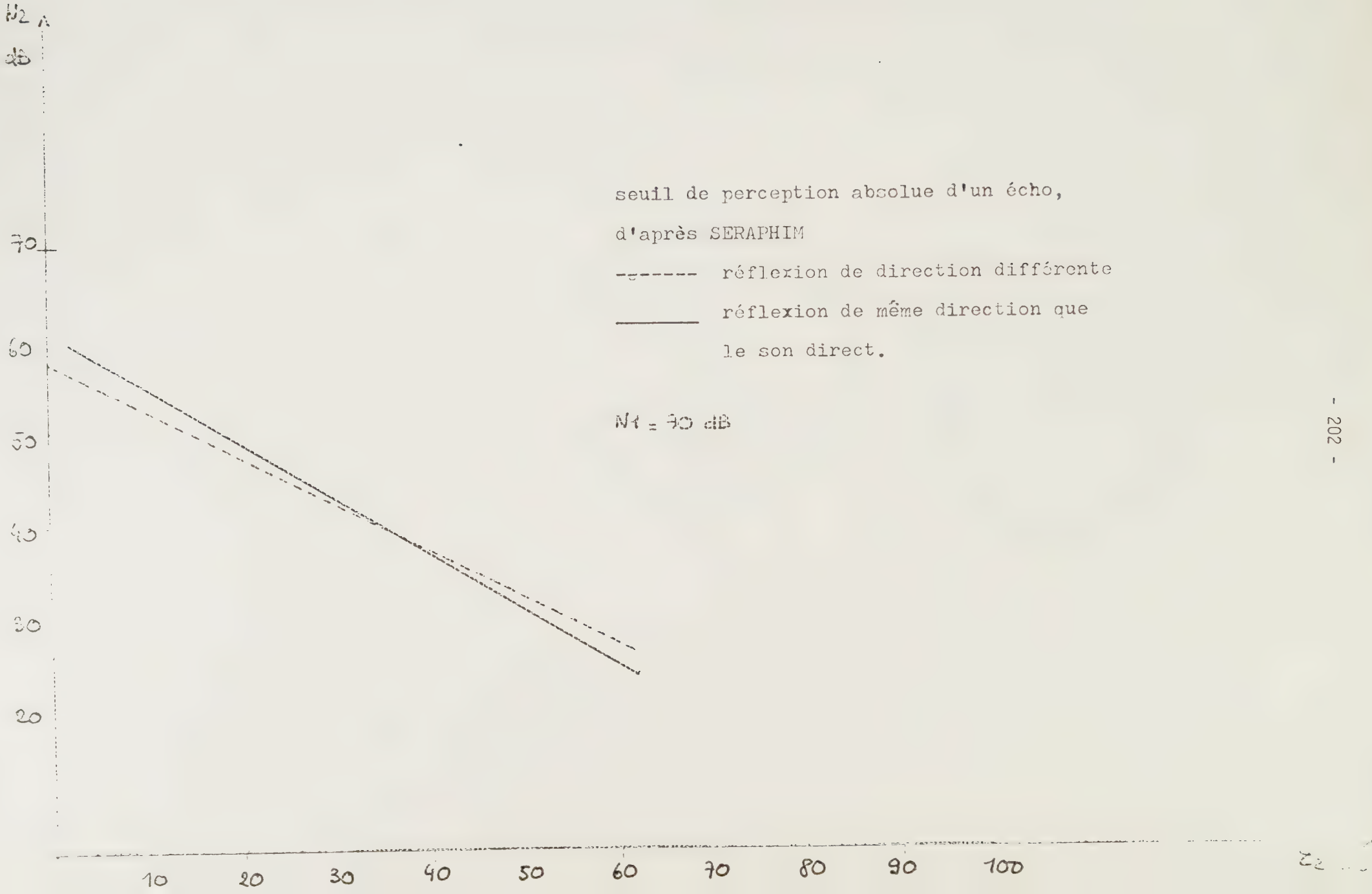
Variation du délai critique d'un écho de même intensité que le son direct, d'après HAAS.

ou $N_{2S}(\varphi_2)$. En ce qui concerne les variations du seuil avec θ_2, Haas a bien constaté comme nous que le seuil, ou le délai critique passe par un minimum lorsque $\theta_2 = 0°$; cependant, il trouve un maximum pour $\theta_2 = 180°$, résultat opposé à celui que nous avons obtenu. Pour $\theta_2 = 0°$, la courbe $\tau_c (\varphi_2)$ passe par un minimum pour $\varphi_2 = 0°$ et par un maximum pour $\varphi_2 = 145°$, ce qui correspond bien à ce que nous avions observé pour les variations de $N_{2S} (\varphi_2)$.

Avant d'essayer de donner une explication à ces résultats, nous mentionnerons ceux obtenus par Seraphim, et relatifs au seuil de perception absolu d'une réflexion (Fig. 5.16). Pour des délais inférieurs à 35 ms environ, Seraphim a trouvé que le seuil de perception absolu était plus élevé lorsque la réflexion avait la même direction que le son direct: ce résultat confirme que le masquage est d'autant plus efficace que la direction du son direct correspond à celle de la réflexion envisagée: autrement dit, on se rend compte très facilement de la présence d'une source secondaire, même de niveau très faible, lorsque cette source a une direction différente de la source primaire. Cependant, le résultat trouvé lorsqu'on s'occupe du seuil de nuisance de l'écho est justement à l'opposé de celui-là. N_{2S} est le plus faible lorsque la direction θ_2 est celle du son direct. C'est là un effet paradoxal, opposé au diagramme de réception de l'oreille (Fig. 5.17), la sonie d'un signal émanant d'une région θ_2 comprise entre 30° et 150° étant plus importante que celle d'un signal qui provient de 0° ou 180°.

Sans avoir la prétention d'en donner une interprétation profonde - remarquons que Haas ne cherche pas à justifier ce résultat - nous dirons que cet effet est dû à l'audition binaurale, et sans doute aux processus cérébraux qui relient les informations reçues par l'une et l'autre des oreilles: si $\theta_1 = \theta_2 = 0°$, l'oreille droite et l'oreille gauche recevront le même signal; par contre, si par exemple $\theta_2 = 30°$, l'oreille gauche recevra la réflexion avec une sonie accentuée, l'oreille droite avec une sonie au contraire diminuée. Peut-on penser, avec Kürer [3.42], que l'auditeur se fie à l'information la plus favorable, en l'occurence celle de l'oreille droite? Ou bien faut-il penser que l'élargissement de la source sonore rend d'auditeur plus tolérant à l'égard de l'effet d'écho?

Toujours est-il que pour $\theta_2 = 180°$, le seuil mesuré est aussi bas que dans le cas où $\theta_2 = 0°$: pour cette situation, les deux oreilles reçoivent également le même signal acoustique.



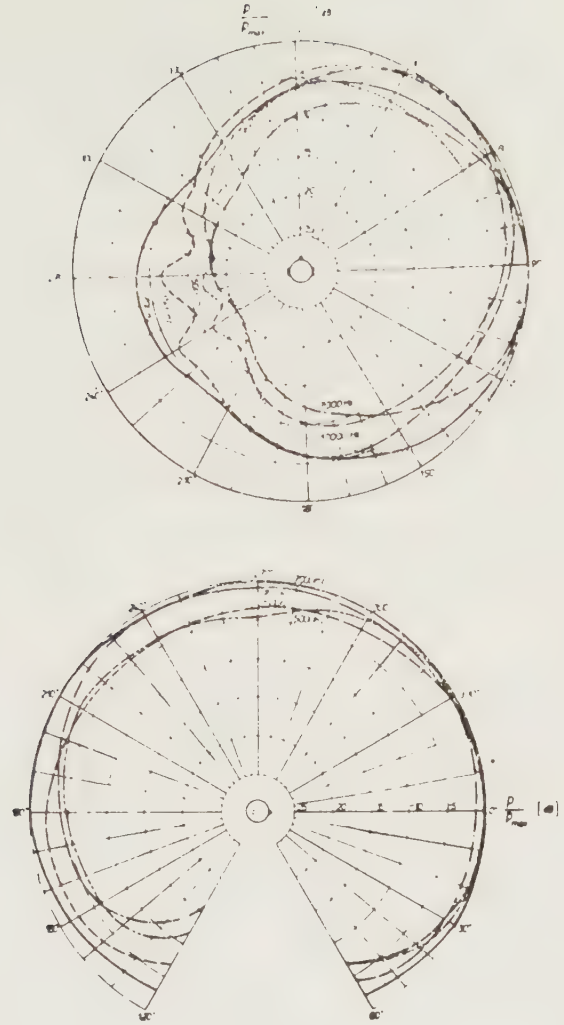
seuil de perception absolue d'un écho,

d'après SERAPHIM

------ réflexion de direction différente

_____ réflexion de même direction que le son direct.

N1 = 70 dB



. : diagramme de réception de l'oreille droite, dans le plan horizontal, et des deux oreilles, dans le plan vertical, d'après Schirmer, cité par Janusha et Spronck] .

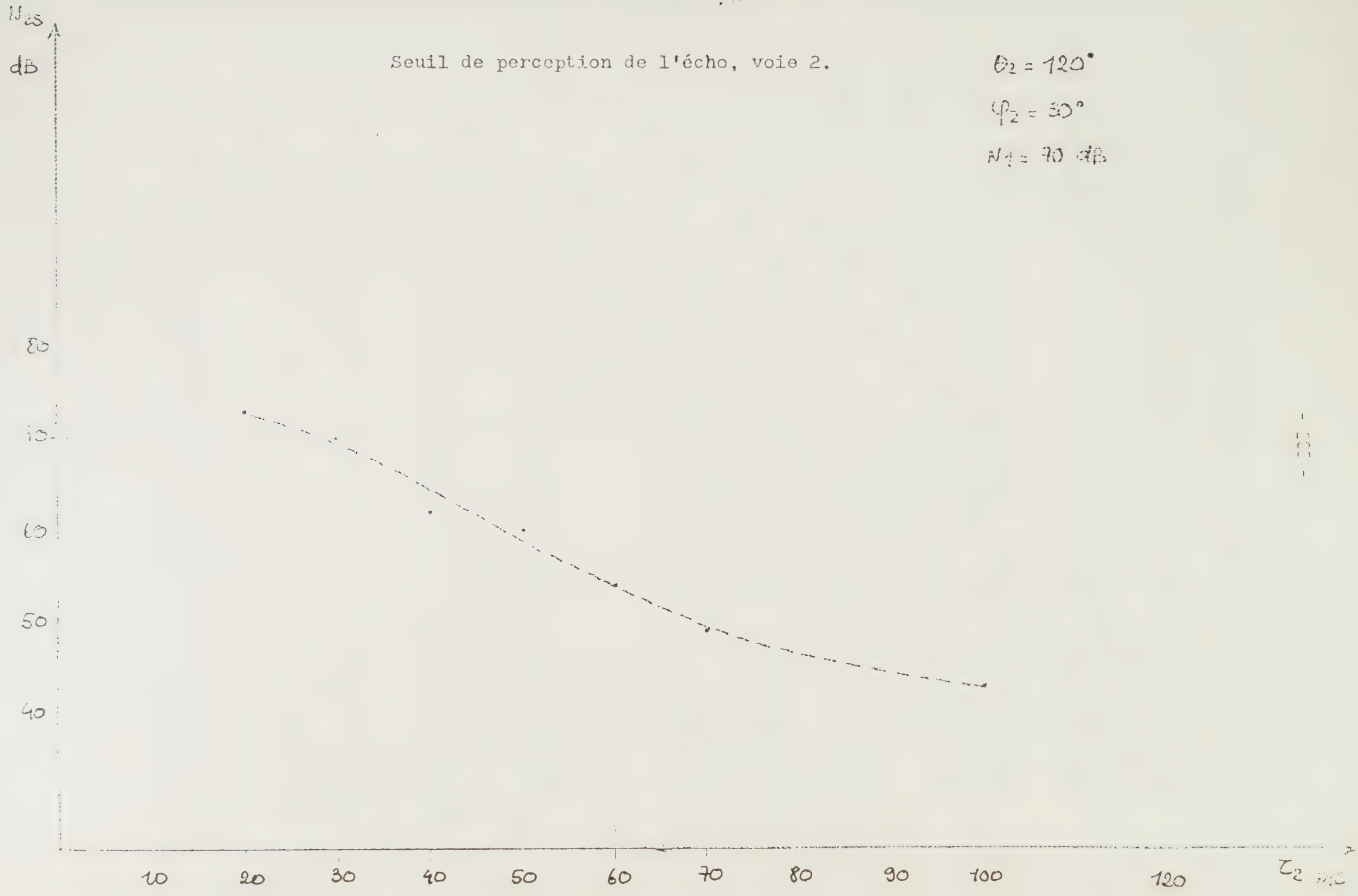
On sait d'autre part que le fait de mélanger deux signaux sonores dont l'un est décalé par rapport à l'autre crée un effet de coloration, par déphasage, souvent appelé "effet de filtre en treillis", et étudié par plusieurs auteurs [5.22] [5.26] [5.27] . Cet effet, perturbateur pour la parole, semble par conséquent très sensible si les directions des haut-parleurs sont voisines; au contraire, il disparaît lorsque ces directions sont écartées, dans le plan horizontal.

Envisageons maintenant ce qui se passe dans le plan vertical: pour toutes les valeurs de φ_2 envisagées, les deux oreilles reçoivent le même signal: l'augmentation du seuil de perception qui apparaît aux alentours de $\varphi_2 = 150°$ ne saurait être justifiée par la différence des signaux reçus à droite et à gauche. Mais le diagramme de réception fait apparaître une baisse de sensibilité dans cette zone, ce qui justifie l'augmentation du seuil de perception.

Pour θ_2 donné, si φ_2 est de l'ordre de 30° à 60° et différent de 0° ou 180°, on a trouvé une légère augmentation du seuil de perception par rapport à la situation $\varphi_2 = 0°$, ainsi qu'on peut en juger sur l'exemple de la figure 5.18. Ce résultat est en bon accord avec ceux de Haas.

Pour terminer ce paragraphe, nous reviendrons sur le cas de l'écho provenant du plan vertical $\theta = 0°$. Kürer [3.42] évoque le cas d'une salle où l'intelligibilité serait nettement plus faible que celle prévue par la méthode du "centre de gravité binaural", satisfaisante dans de nombreux cas. Or, cette salle a des murs absorbants, alors que le plafond est réfléchissant. Nos mesures justifient ce résultat: une réflexion, même proche, risque d'être perturbatrice; il semble par conséquent que cette disposition d'absorbants soit à éviter.

La situation que nous avons envisagée jusqu'à présent, c'est-à-dire celle de la première réflexion qui suit le son direct, est applicable à tous les échogrammes. Dès qu'au son direct, nous supposerons associées certaines réflexions, nous serons obligés de n'envisager que des cas particuliers; la situation deviendra différente dès qu'on changera l'un des quatre paramètres qui caractérisent une réflexion, à savoir son niveau, son retard, son site et son azimut. Nous tâcherons de considérer les situations les plus caractéristiques.


Seuil de perception de l'écho, voie 2.
$\theta_2 = 120°$
$\varphi_2 = 30°$
$N_1 = 70$ dB
N_{2S}
dB
80
70
60
50
40
10
20
30
40
50
60
70
80
90
100
120
τ_2 ms

d) - Cas d'une source primaire multidirectionnelle.-

Nous allons maintenant envisager un certain nombre d'exemples: le son direct peut être suivi de réflexions très proches qui lui sont parfaitement intégrées. Ce sera le cas d'une salle de concerts où la scène est entourée d'une "coquille", destinée justement à jouer ce rôle; quel est alors le seuil de perception d'un écho qui arriverait après, donné par le plafond ou les murs latéraux?

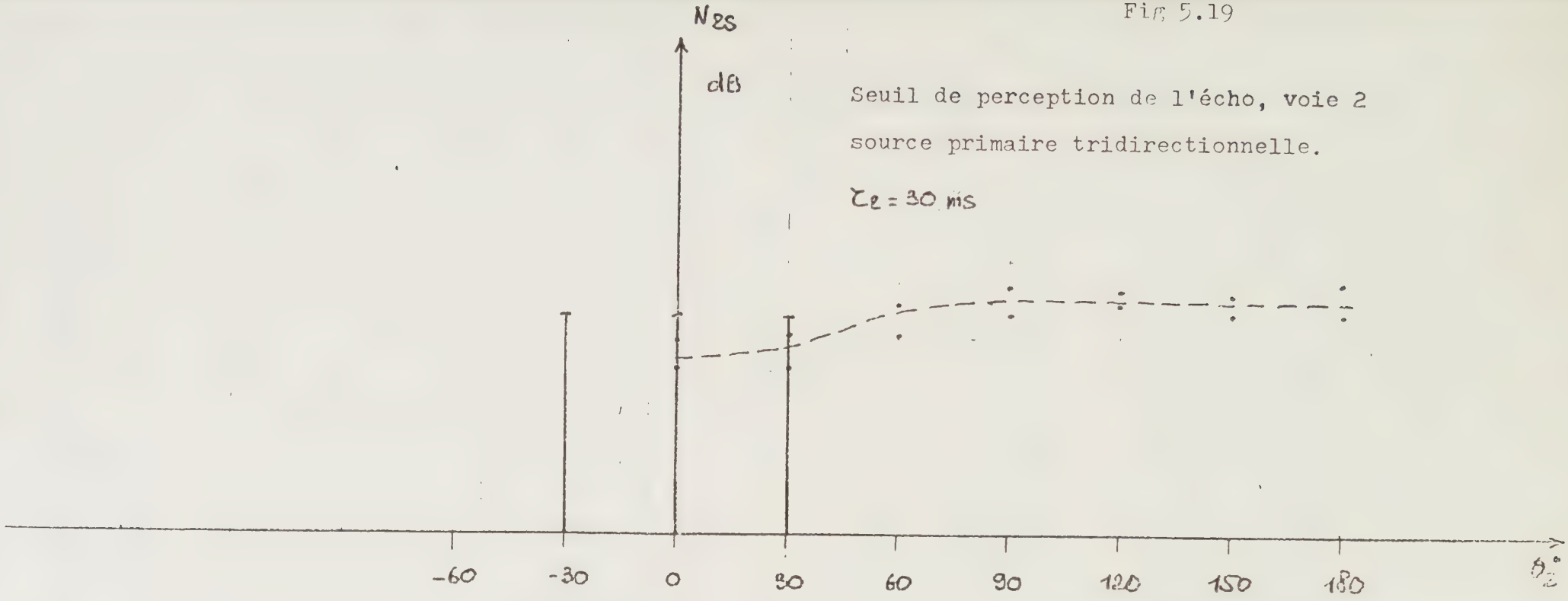
Comme le dispositif expérimental ne nous permettait pas d'obtenir des délais très faibles, nous nous sommes contentés de mettre en parallèle sur la voie 1 deux coffrets de haut-parleurs qui simulent des réflexions très proches. Comme il semblait que les variations du seuil avec l'azimut soient plus caractéristiques que celles relevées dans le plan vertical, nous avons travaillé avec des angles φ toujours nuls. Nous appellerons "source primaire" l'ensemble des coffrets alimentés par la voie 1, et dont les directions sont θ_{11} , θ_{12} et θ_{13}. θ_{11} sera toujours nul, θ_{12} et θ_{13} auront diverses valeurs. Les courbes N_{2S} (τ_2) ont des pentes très voisines quelles que soient les valeurs de θ_{12} et θ_{13} , aussi nous considérerons le seuil N_{2S} à 30 ms comme caractéristique de la situation de ces courbes. L'ensemble des directions θ_1 créent une "source fantôme" de situation moyenne, et qui peut par exemple être concentrée en $\theta = 0°$.

α) Source primaire symétrique.

Si la source primaire est étroite, par exemple pour $\theta_{12} = 30°$ et $\theta_{13} = -30°$, le seuil N_{2S} à 30 ms est minimum pour $\theta_2 = 0°$, ce qui correspond à la situation de θ_{11} et de la source fantôme résultante (Fig. 5.19). L'écart ΔN_{2S} entre le minimum et le maximum de N_{2S} est de l'ordre de 5 dB. Il faut noter que le seuil moyen est assez faible, et sensiblement inférieur au niveau total $N_1 \simeq 65$ dB, chaque direction θ_{11} , θ_{12} et θ_{13} ayant un niveau de 60 dB.
D'autre part, le minimum observé précédemment pour $\theta_2 = 180°$ disparaît.
Si la source s'élargit ($\theta_{12} = 60°$, $\theta_{13} = -60°$), le minimum de N_{2S} , toujours situé dans la région de la source primaire, est moins marqué: le seuil moyen est plus élevé, alors que l'écart entre le maximum et le minimum de N_{2S} est de l'ordre de 4 dB.

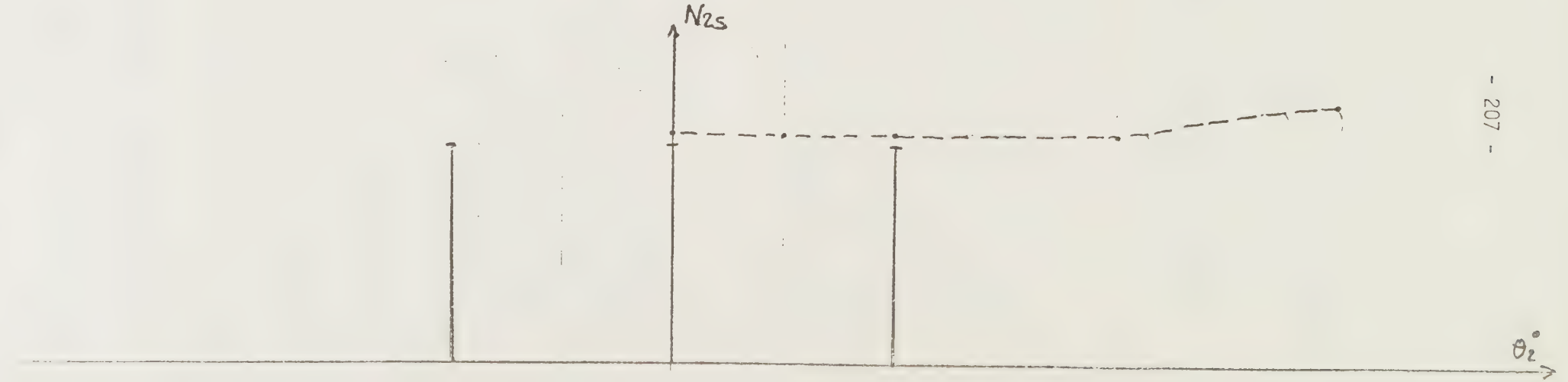
Si l'on envisage les directions $\theta_{12} = 90°$ et $\theta_{13} = -90°$, on observe un effet de masquage direct pour les directions $\theta_2 = 0°$ et $\theta_2 = \pm 90°$.

Fig 5.19



Seuil de perception de l'écho, voie 2
source primaire tridirectionnelle.

$\tau_2 = 30$ ms



Au contraire, le seuil est minimum pour $\theta_2 = \pm 30°$ (Fig. 5.20). L'écart ΔN_{2S} est assez faible.

Pour la situation $\theta_{12} = 120°$, $\theta_{13} = -120°$, le minimum de N_{2S} pour $\theta_2 = 0°$ est plus marqué: il existe un autre minimum pour $\theta_2 = 180°$, correspondant à la situation de la source fantôme créée par les directions θ_{12} et θ_{13}.

Au contraire, pour $\theta_2 = \pm 120°$, on a affaire à un effet de masquage direct, le seuil du son masqué est relativement élevé pour une direction correspondant à celle du son masquant.

β) Source primaire dissymétrique et bilatérale. (Fig. 5.21 et 5.22)

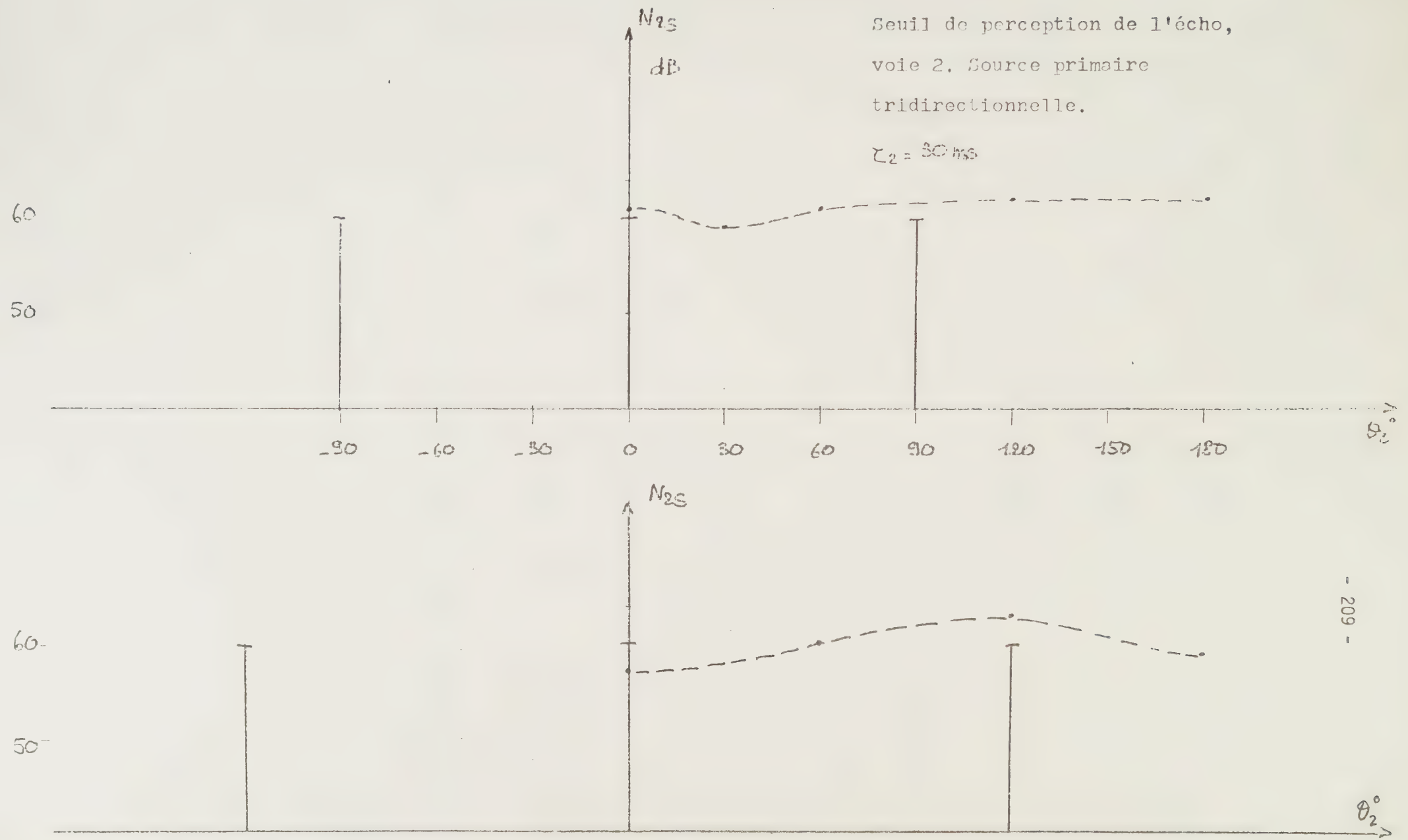
Envisageons par exemple la situation $\theta_{12} = -30°$, $\theta_{13} = 60°$. Le minimum de N_{2S} correspond à la direction de la source primaire, avec un décalage sensible du côté de la source fantôme résultante. L'écart de variation de N_{2S} est de l'ordre de 10 dB. On trouve un résultat assez analogue pour $\theta_{12} = -60°$ et $\theta_{13} = 90°$. Si l'on envisage une source nettement plus large, soit $\theta_{12} = -60°$ et $\theta_{13} = 120°$, on observe un minimum de N_{2S} dans la région $\theta_2 = 0$, alors que pour $\theta_2 = 0_{12}$ ou $\theta_2 = 0_{13}$ apparaît un effet de masquage direct, entraînant un maximum de N_{2S} .

γ) Cas d'une source monolatérale. (Fig. 5.23 et 5.24)

Supposons que les réflexions proches proviennent d'un seul côté, ce qui serait le cas pour un auditeur assis dans une salle, à proximité d'un mur latéral. Pour les angles $\theta_{12} = 30°$ et $\theta_{13} = 60°$ ou $\theta_{12} = 60°$ et $\theta_{13} = 120°$ le seuil N_{2S} est minimum lorsque θ_2 est situé dans la région de la source primaire: ce minimum ne correspond pas à $\theta_2 = 0$, mais il est décalé vers le centre de la source primaire. Le seuil est maximum dans la région opposée à la source de centre θ_1, soit θ_1 -180°.
Les variations du seuil sont de l'ordre de 10 dB. C'est le même effet qu'on relève si la source est encore plus large ($\theta_{12} = 90°$, $\theta_{13} = 180°$), mais lorsque θ_2 passe par la valeur intermédiaire $\theta_{12} = 90°$, on observe nettement une augmentation locale du seuil N_{2S}, dû à un effet de masquage direct.

De l'ensemble de ces mesures, nous tirerons les conclusions suivantes :

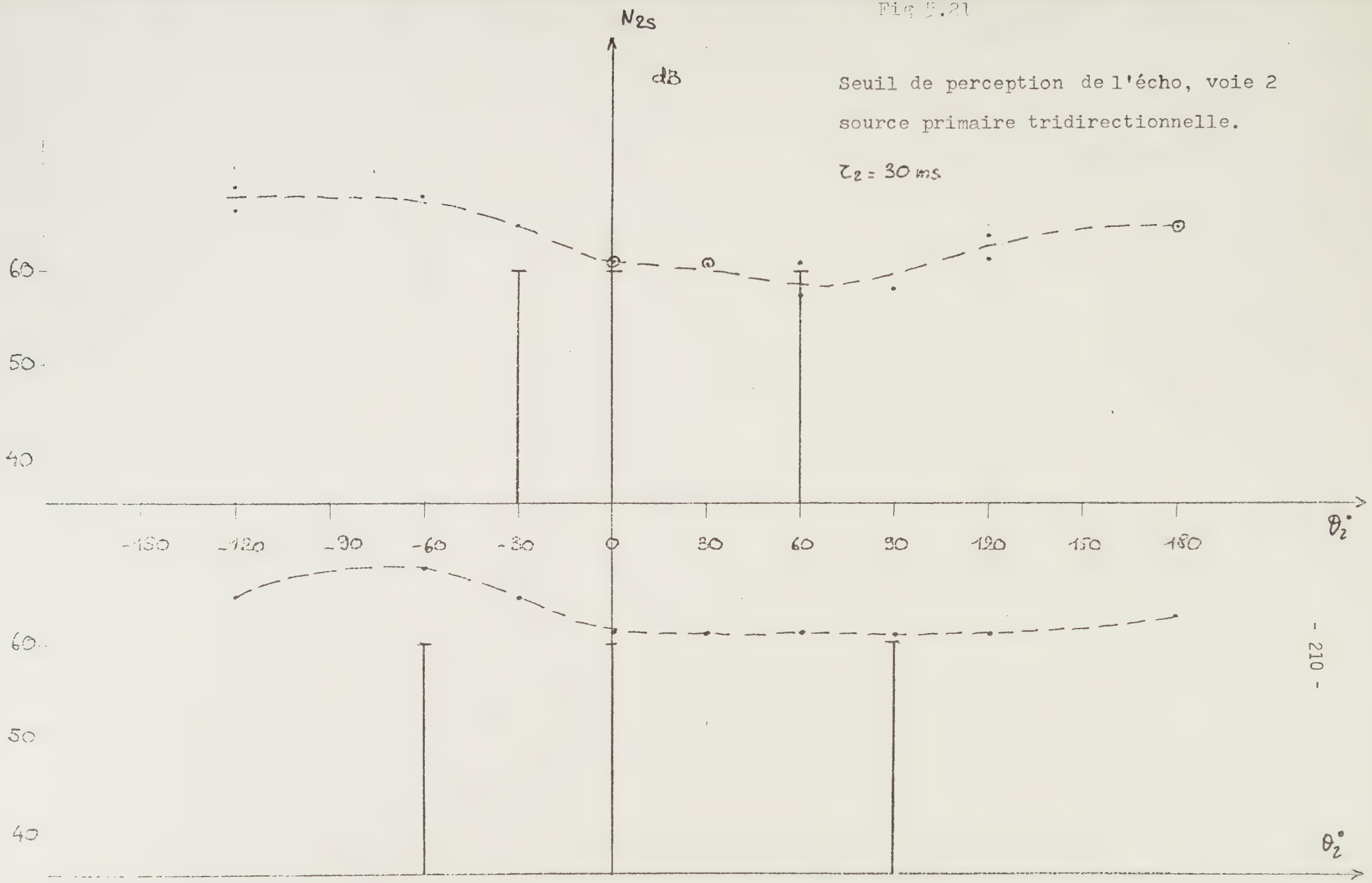
- d'une façon générale, le seuil moyen $\overline{N}_{2S}$ est d'autant plus élevé que la source primaire est plus large.
- si la source primaire a une largeur réduite (écart entre les directions θ_1



Seuil de perception de l'écho, voie 2. Source primaire tridirectionnelle.

$\tau_2 = 30$ ms

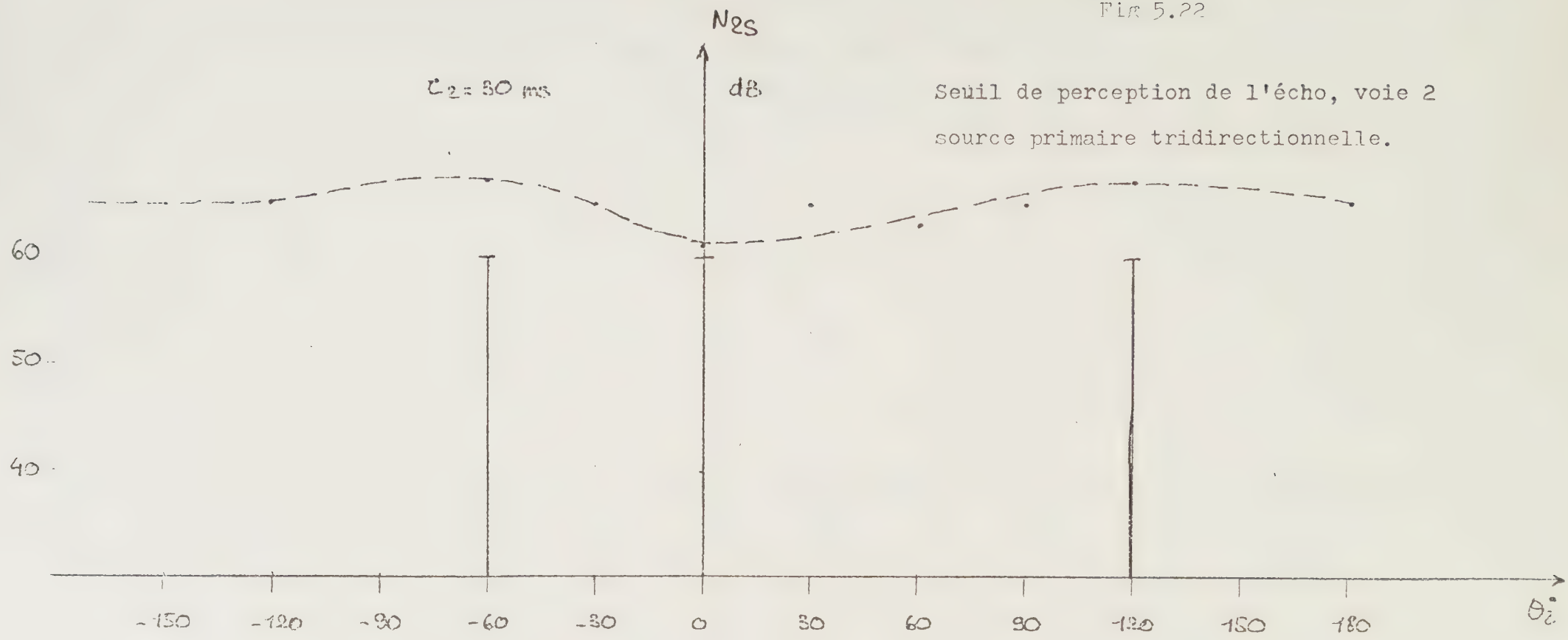
Fig 5.21



Seuil de perception de l'écho, voie 2
source primaire tridirectionnelle.

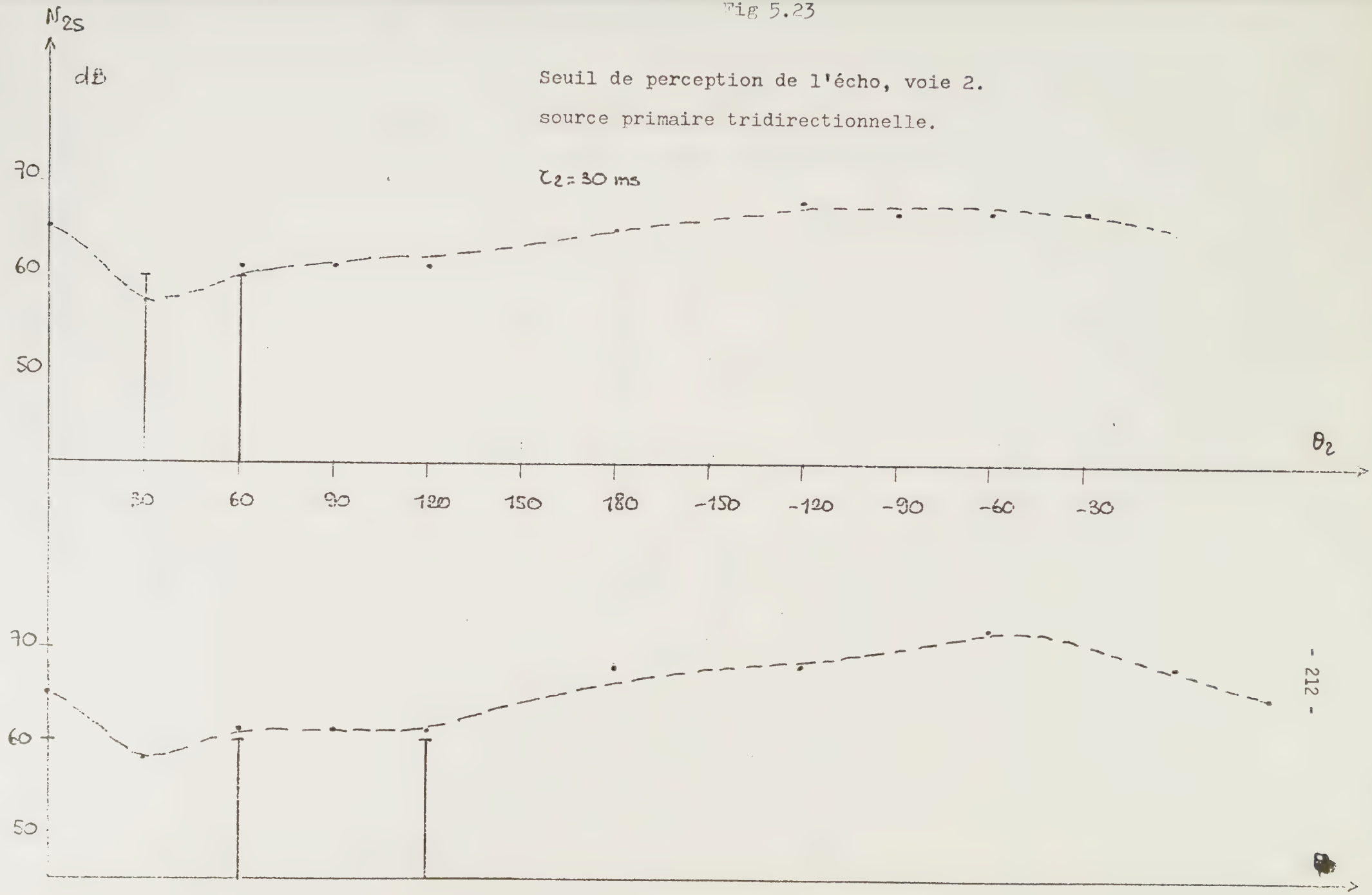
$\tau_2 = 30$ ms

Fig 5.22



Seuil de perception de l'écho, voie 2
source primaire tridirectionnelle.

Fig 5.23

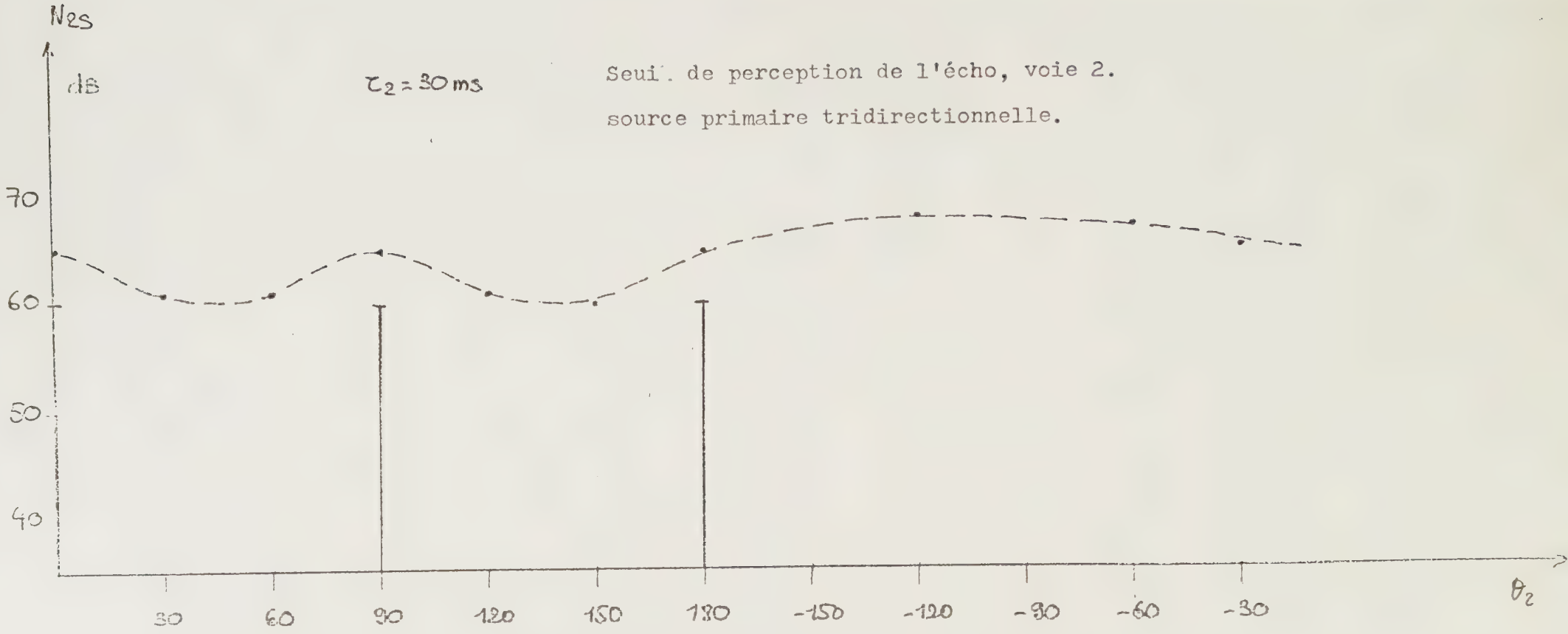


Seuil de perception de l'écho, voie 2.

source primaire tridirectionnelle.

$\tau_2 = 30$ ms

Fig I.24



Seuil de perception de l'écho, voie 2.

source primaire tridirectionnelle.

extrêmes inférieur à 180°), on observe un minimum du seuil N_{2S} lorsque θ_2 est dans le domaine de la source primaire, et un maximum dans la direction opposée.

- au contraire, si la source est large (écart des directions θ_1 supérieur à 180°), il peut apparaître, dans certaines conditions, un effet de masquage direct pour $\theta_2 = 0$ et plus généralement pour les autres directions θ_{12}, θ_{13} de la source primaire.

La Fig. 5.25 résume les situations extrêmes des courbes N_{2S} (τ_2) obtenues lorsque la source primaire est réduite au son direct.

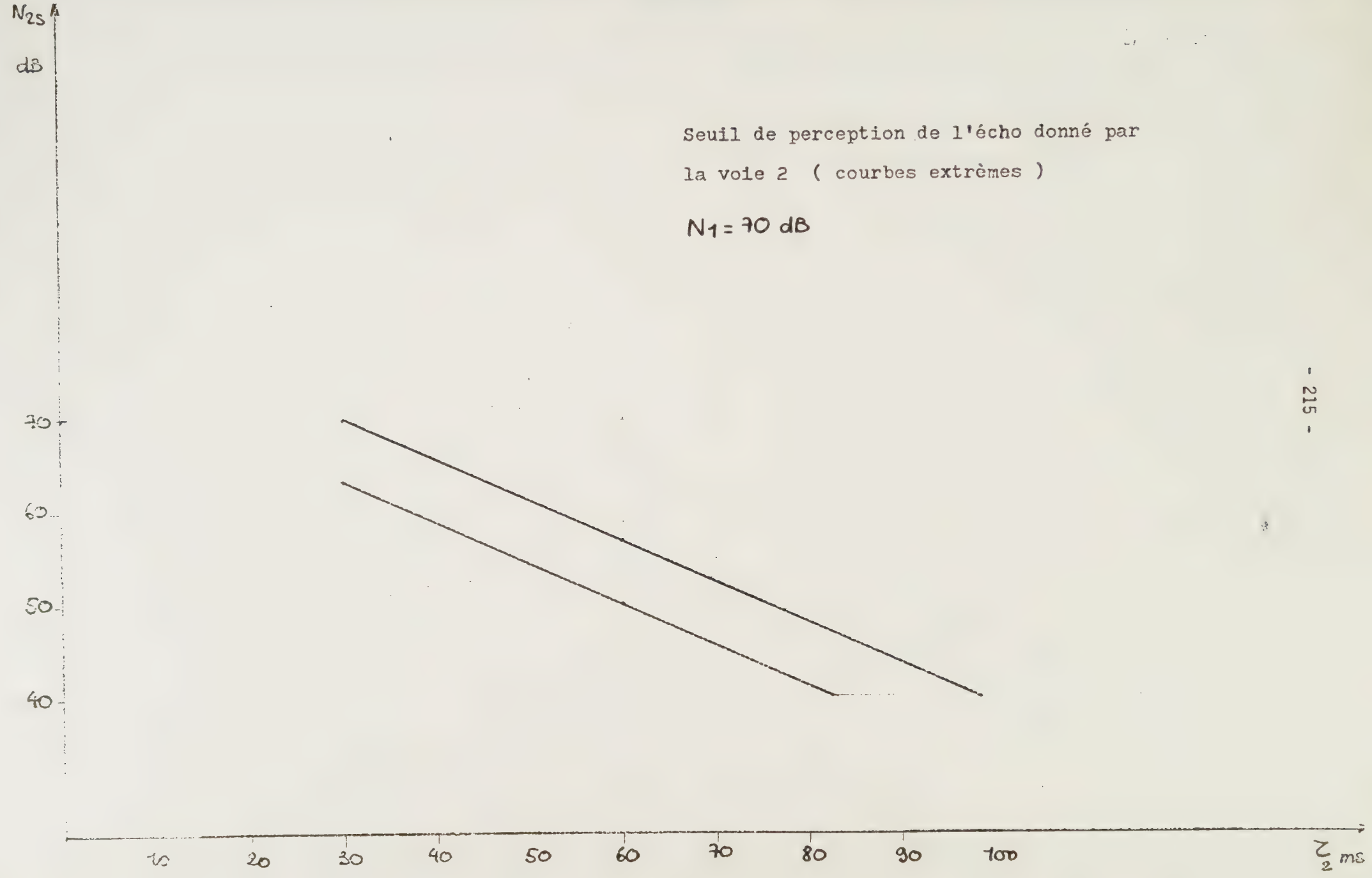
5.5.3.- Seuil de perception de l'écho, voie 3.-

La source primaire est constituée par le son direct et une première réflexion de même niveau, qui lui est associée : d'après la figure 5.12, cette réflexion, retardée de 15 ms ou 30 ms, doit avoir une direction θ comprise entre 30° et 150°.

On remarquera que la source primaire ainsi constituée est obligatoirement monolatérale. Les courbes N_{3S} (τ_3) apparaissent comme nettement décalées vers la droite par rapport à celles obtenues dans le cas de l'écho donné par la voie 2. La figure 5.26 donne les situations extrêmes dans le cas où la première réflexion (voie 2) est à 15 ms ou à 30 ms, suivant les directions des voies 2 et 3. En comparant ces courbes à celles de la figure 5.25, on voit que le seuil, pour une valeur de τ donné, est relevé d'une dizaine de décibels. Par conséquent, les réflexions qui suivent le son direct et lui sont intégrées modifient très sensiblement le seuil de perception de l'écho suivant. Cette constatation, qui peut paraître évidente, amène à revoir les courbes d'intégration auditive obtenues et utilisées par des auteurs tels que Niese ou Lochner : celles-ci ne mettent en cause que le son direct et une seule réflexion, ce qui revient en fait à limiter le seuil de perception d'un écho aux courbes de la Fig. 5.25.

Bien entendu, la présence de courbes multiples, suivant les différents paramètres des réflexions associées au son direct n'est pas faite pour simplifier le problème, mais nous ne pensons pas qu'on doive négliger la complexité de la situation. Nous donnerons deux exemples de courbes N_{3S} (τ_3) parmi celles que nous avons tracées : la figure 5.27 donne les courbes obtenues pour $\theta_2 = 30°$ et $\theta_3 = 0°$.

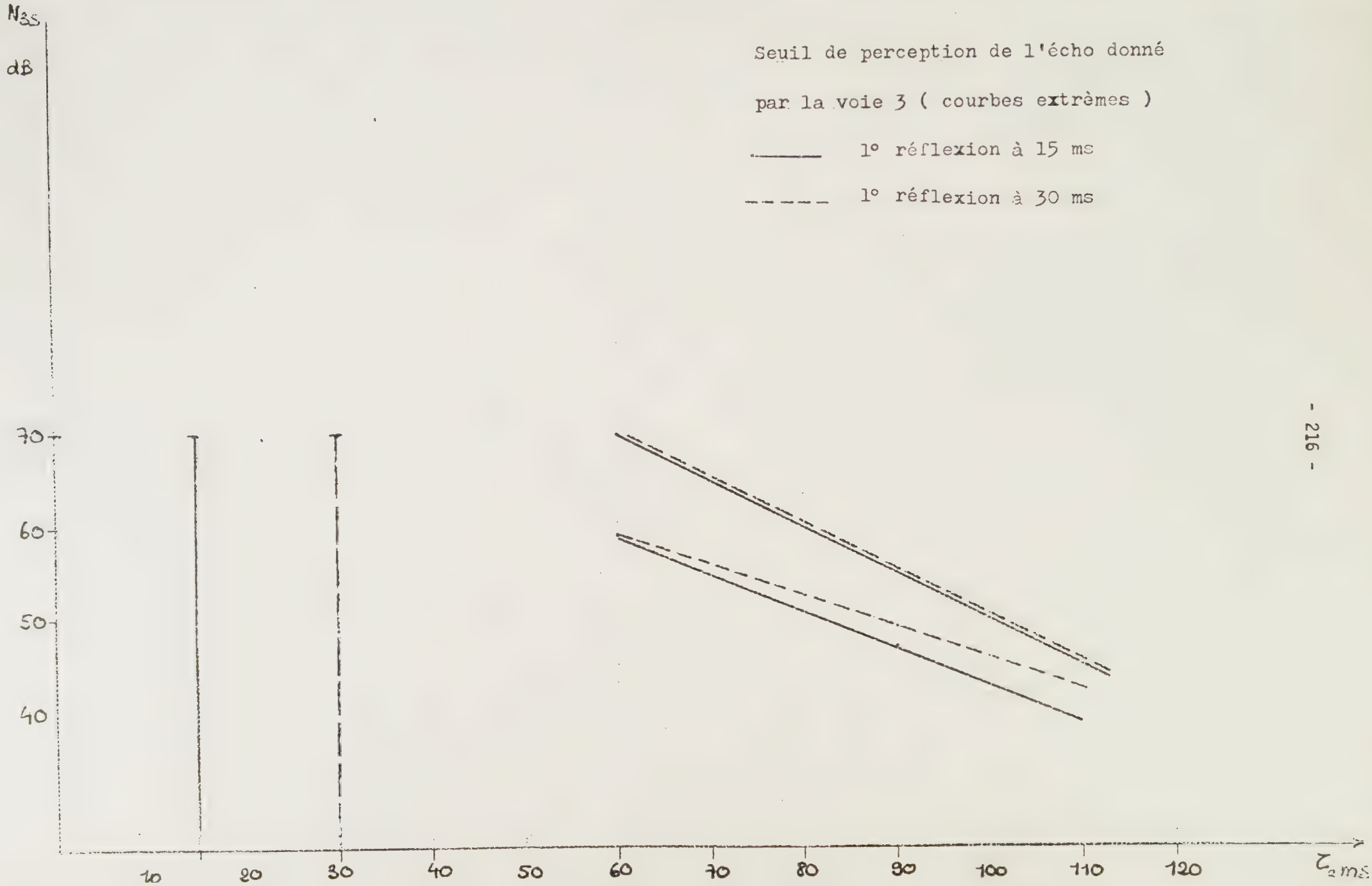
Fig 5.25

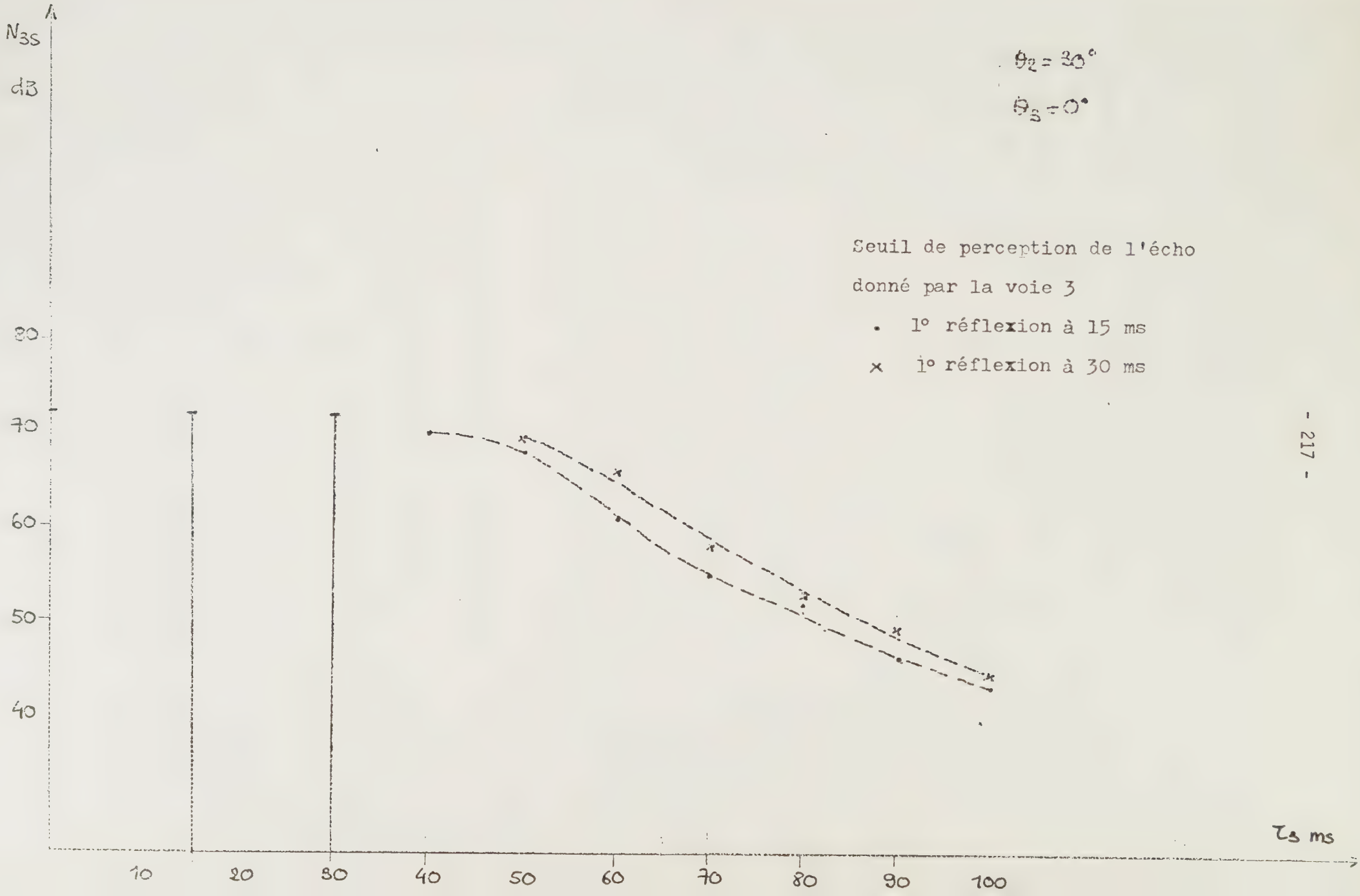


Seuil de perception de l'écho donné par la voie 2 (courbes extrèmes)

$N_1 = 70$ dB

Fig 5.26


N_{ss}
dB
Seuil de perception de l'écho donné
par la voie 3 (courbes extrèmes)
1° réflexion à 15 ms
1° réflexion à 30 ms
70
60
50
40
10
20
30
40
50
60
70
80
90
100
110
120
τ_2 ms


N_{3S}
dB
$\theta_2 = 30°$
$\theta_S = 0°$
Seuil de perception de l'écho
donné par la voie 3
• 1° réflexion à 15 ms
× 1° réflexion à 30 ms
80
70
60
50
40
τ_S ms
10
20
30
40
50
60
70
80
90
100

Les courbes qui correspondent à des retards de 15 ms et 30 ms pour la première réflexion sont assez proches, leur écart est de l'ordre de 2dB seulement : leur pente est de 0,5 dB/ms, leur allure semblable à celle des courbes N_{2S} (τ_2). La figure 5.28 donne les courbes obtenues dans la situation $\theta_2 = 60°$ et $\theta_3 = -90°$. Pour ces valeurs de l'angle θ, le seuil relevé est nettement plus haut que dans la situation précédente, tout au moins lorsque $\tau_2 = 30$ ms. On remarque que les deux courbes sont plus écartées que dans le cas précédent. D'autre part, leur pente est plus forte. Il apparaît ainsi que le seuil de perception de l'écho donné par la voie 3 est d'autant plus élevé que la réflexion associée au son direct est plus retardée. Cet effet est très sensible si la direction de θ_3 est opposée à celle de la source primaire. Les figures 5.29 à 5.32 donnent quelques exemples de courbes N_{3S} (θ_3) obtenues pour $\tau_3 = 60$ ms. On remarque que l'écart de variation de N_{3S} est de l'ordre de 10 dB , et que le maximum du seuil de perception correspond toujours à une direction θ_3 opposée à celle de la source primaire.

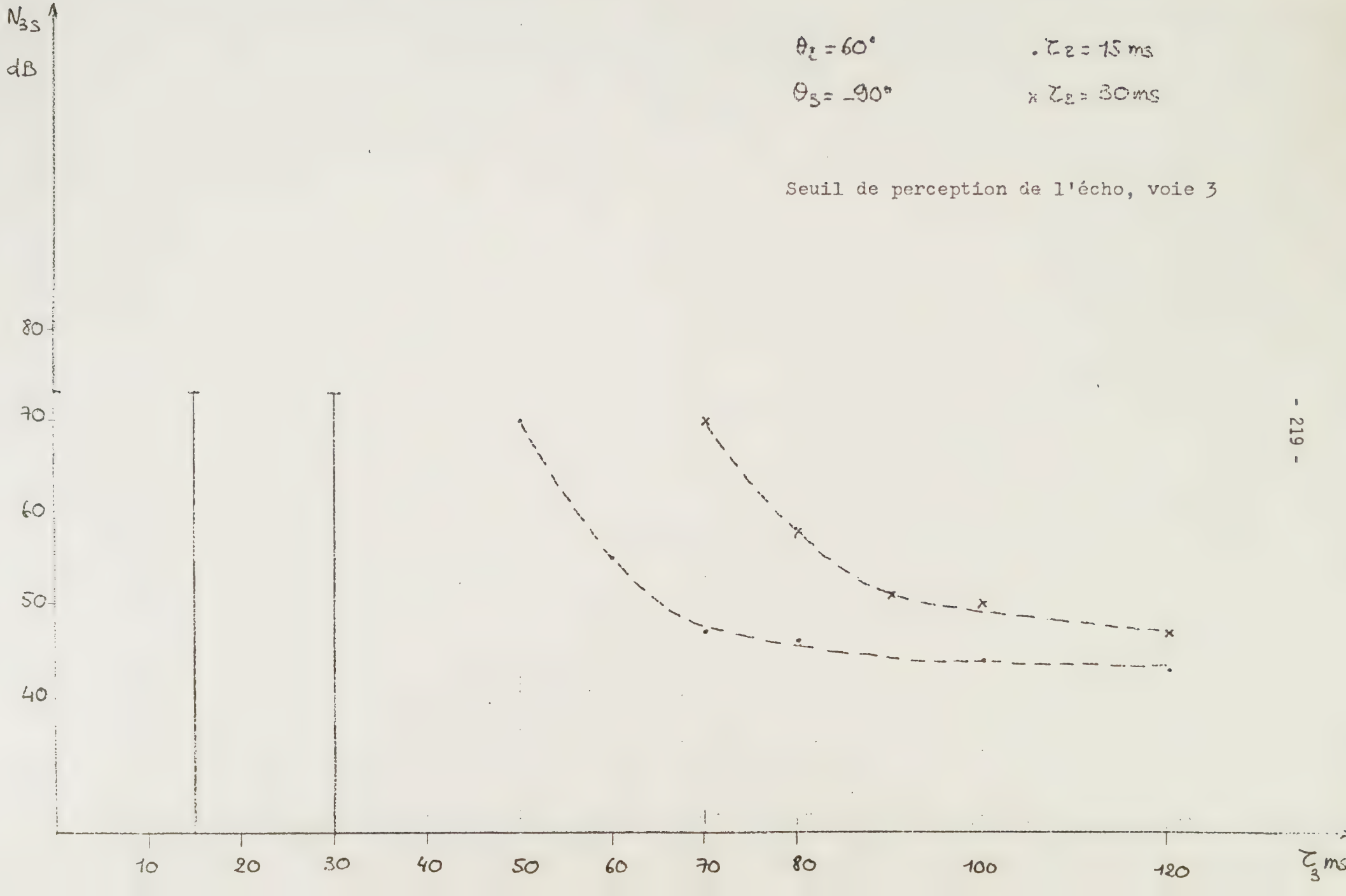
La figure 5.33 donne un exemple de résultats obtenus lorsque la première réflexion a une intensité plus faible que le son direct, soit 3 dB de moins. La courbe N_{3S} (θ_3) a une allure très semblable à celle de la figure 5.30, le seuil à 60 ms relevé pour $\theta_3 \in (0°, 60°)$ est égal à N_1 - 16 dB alors qu'il était égal à N_1 - 10 dB dans le cas de la figure 5.30 : si la source primaire est réduite au son direct, le seuil N_{2S} relevé à 60 ms a pour valeur N - 21 dB (Fig. 5.10), pour $\theta_2 = 0°$. On voit bien quel rôle peut jouer une réflexion située entre l'écho soumis à la mesure, et le son direct.

5.5.4.- Seuil de perception de l'écho, voie 4.-

- a) Cas où les voies 2 et 3 sont monodirectionnelles.-

La source primaire est constituée par le son direct et les voies 2 et 3. Les figures 5.34 et 5.35 donnent deux exemples de courbes N_{4S} (τ_4) relevées dans le cas où $\theta_2 = -30°$ et $\theta_3 = 60°$ avec $\tau_2 = 15$ ms et $\tau_3 = 30$ ms, pour les valeurs $\theta_4 = 0°$ et $\theta_4 = 120°$.
La figure 5.36 donne les situations extrêmes des courbes relevées pour différentes valeurs des angles θ_2, θ_3 et θ_4. Par comparaison avec la courbe de la figure 5.24, courbe globale obtenue pour l'écho donné par la voie 3, il apparaît que la différence est assez faible avec les courbes relatives à $\tau_2 = 30$ ms : il semble donc que c'est la dernière réflexion intégrée au

Fig 5.28



Seuil de perception de l'écho, voie 3

Fig 5.29

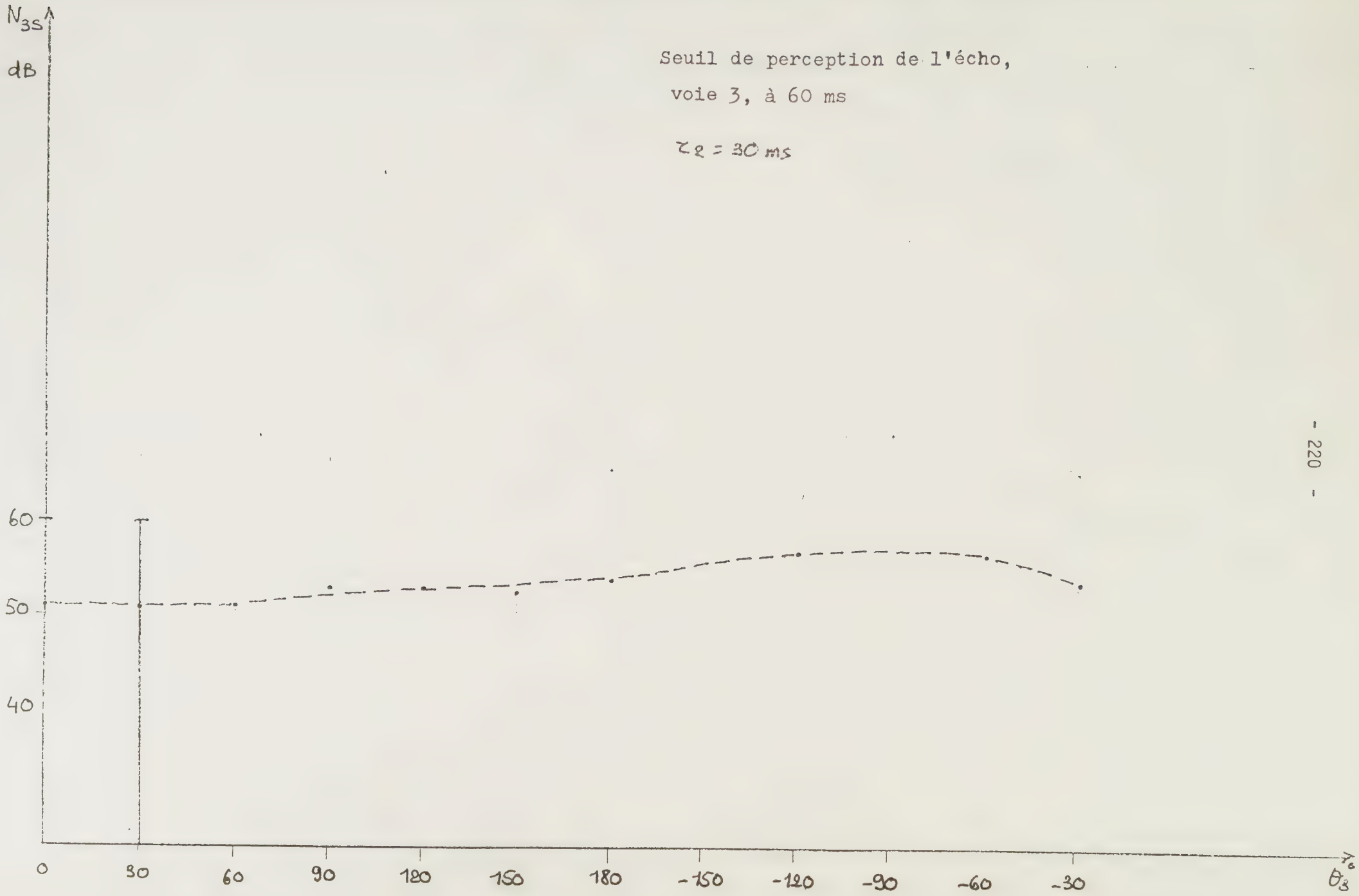

N_{3S}
dB
Seuil de perception de l'écho,
voie 3, à 60 ms
τ_e = 30 ms
60
50
40
0
30
60
90
120
150
180
-150
-120
-90
-60
-30
θ_3

Fig 5.30

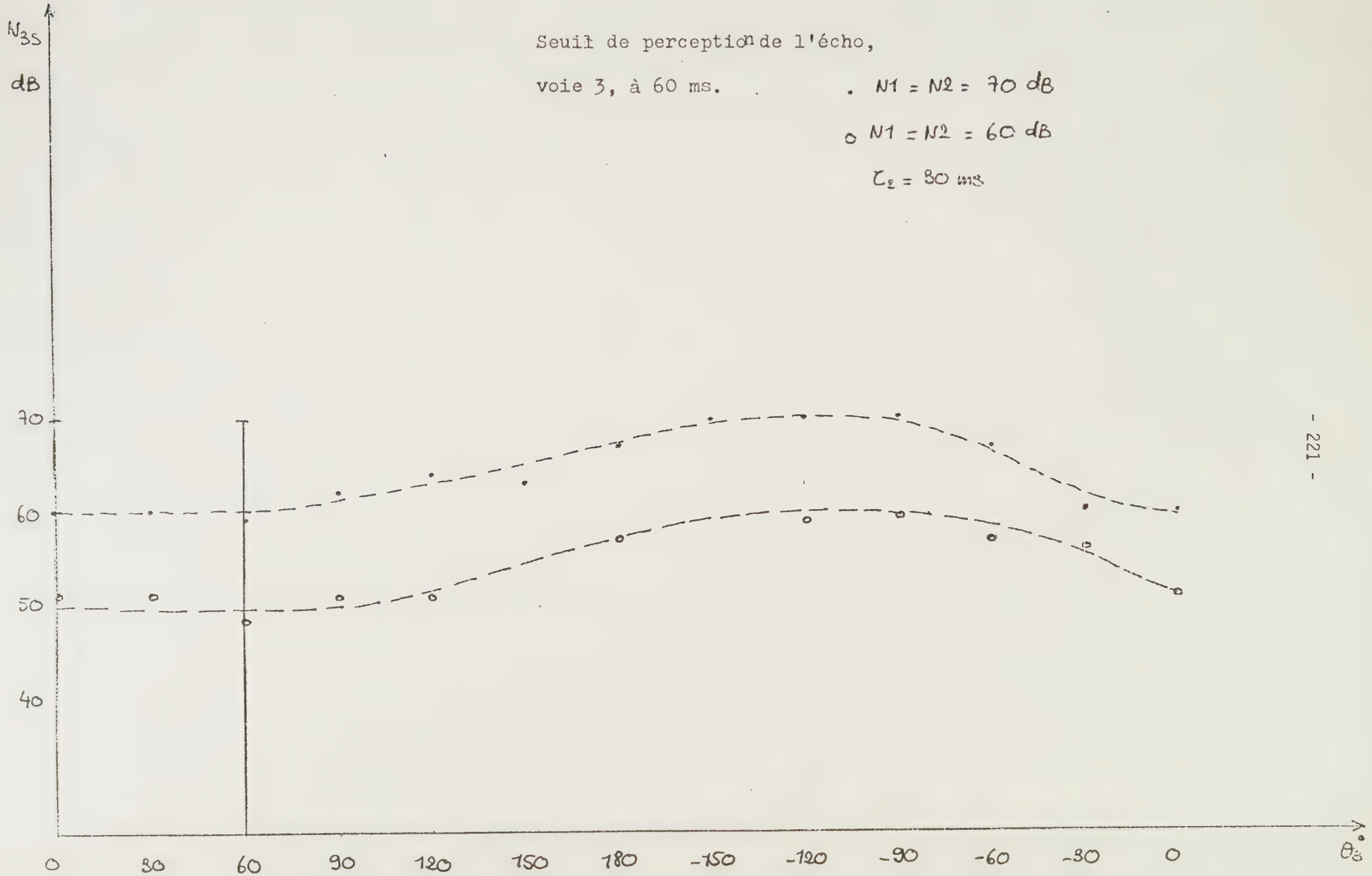
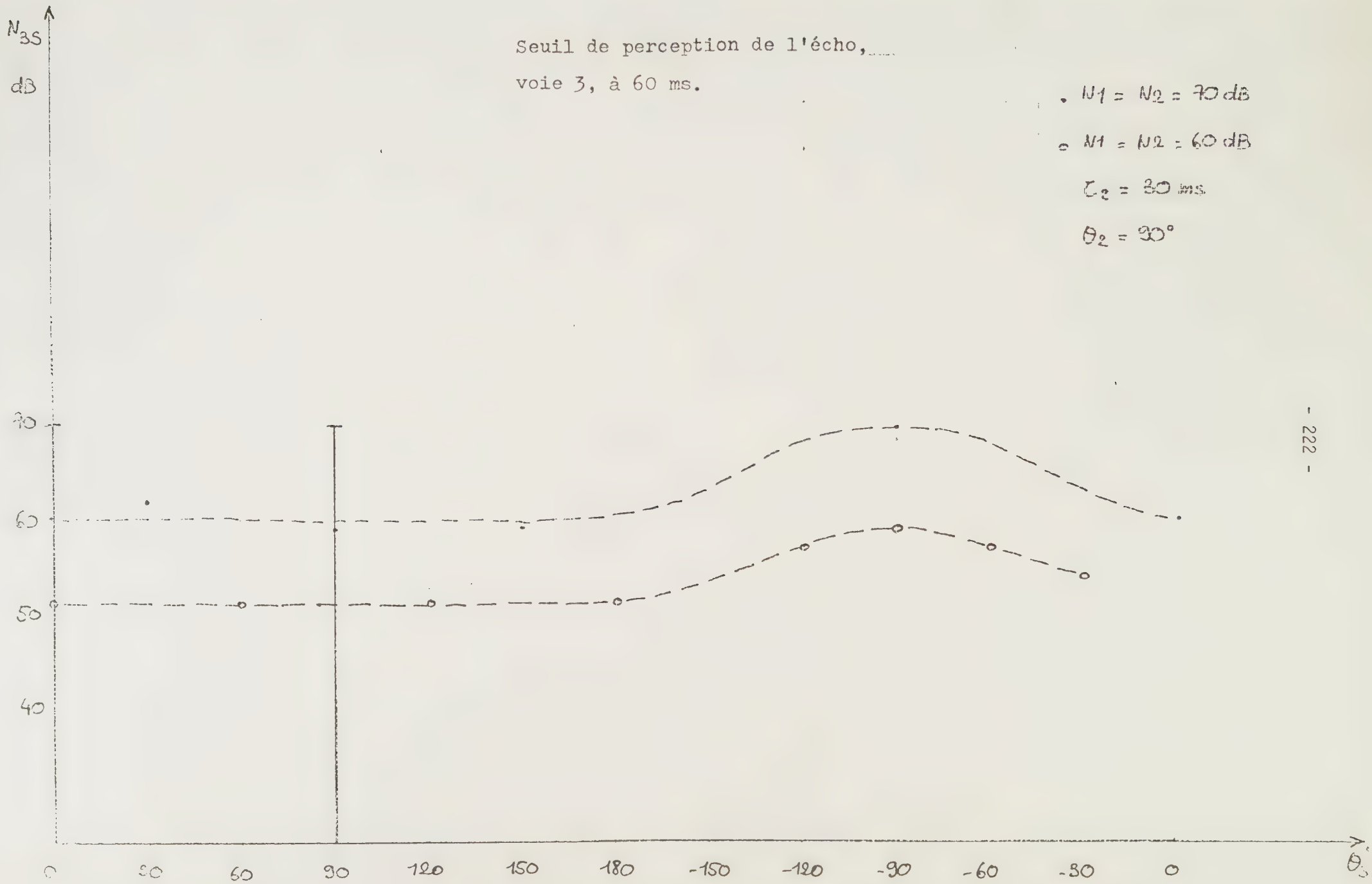

Seuil de perception de l'écho,
voie 3, à 60 ms.
. N1 = N2 = 70 dB
o N1 = N2 = 60 dB
τ_E = 80 ms
N_{3S}
dB
70
60
50
40
0
30
60
90
120
150
180
-150
-120
-90
-60
-30
0
θ_S

Fig 5.31

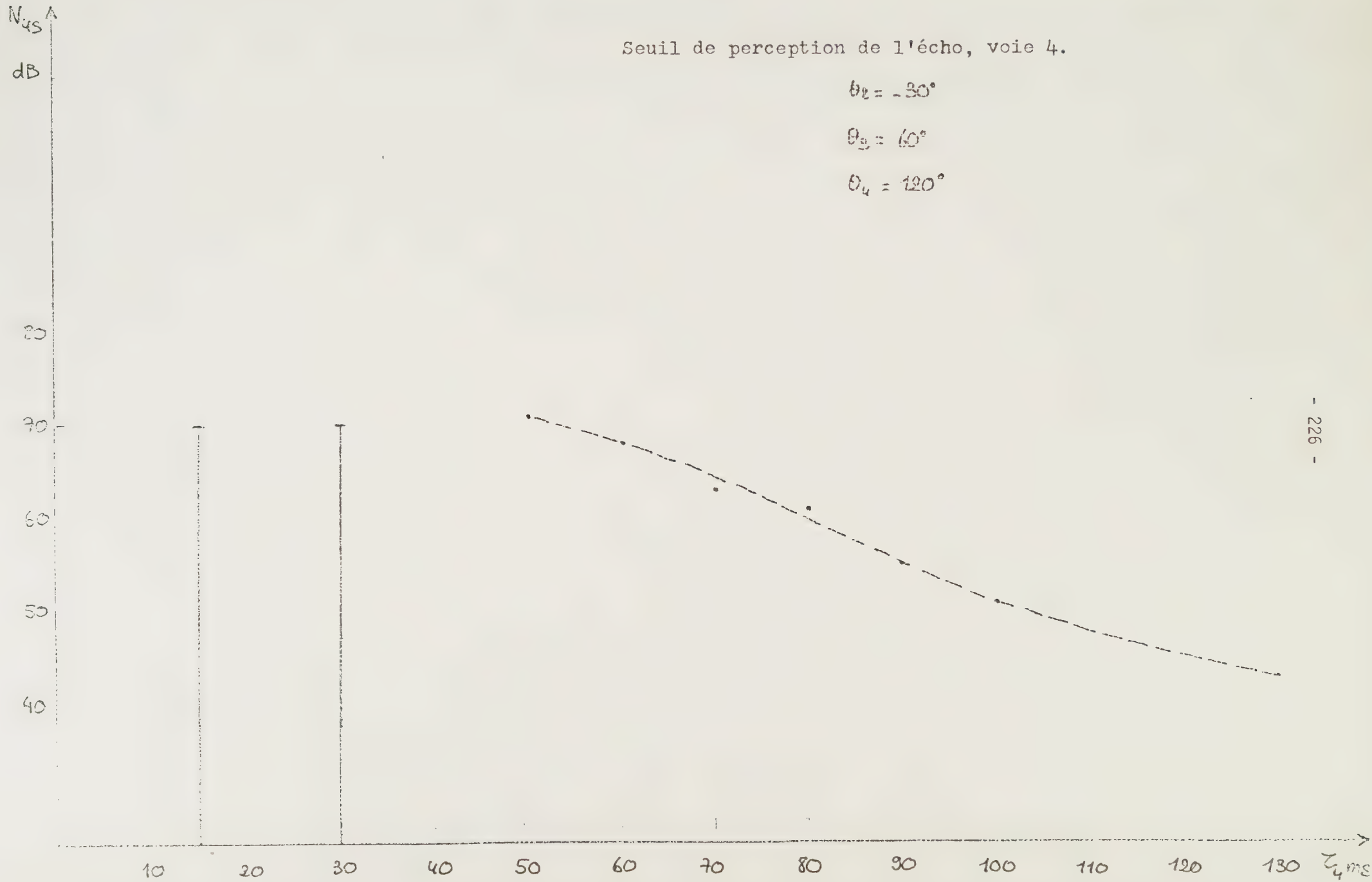

Seuil de perception de l'écho,
voie 3, à 60 ms.
. N1 = N2 = 70 dB
o N1 = N2 = 60 dB
τ2 = 30 ms
θ2 = 30°
N3S
dB
70
60
50
40
0
30
60
90
120
150
180
-150
-120
-90
-60
-30
0
θ3

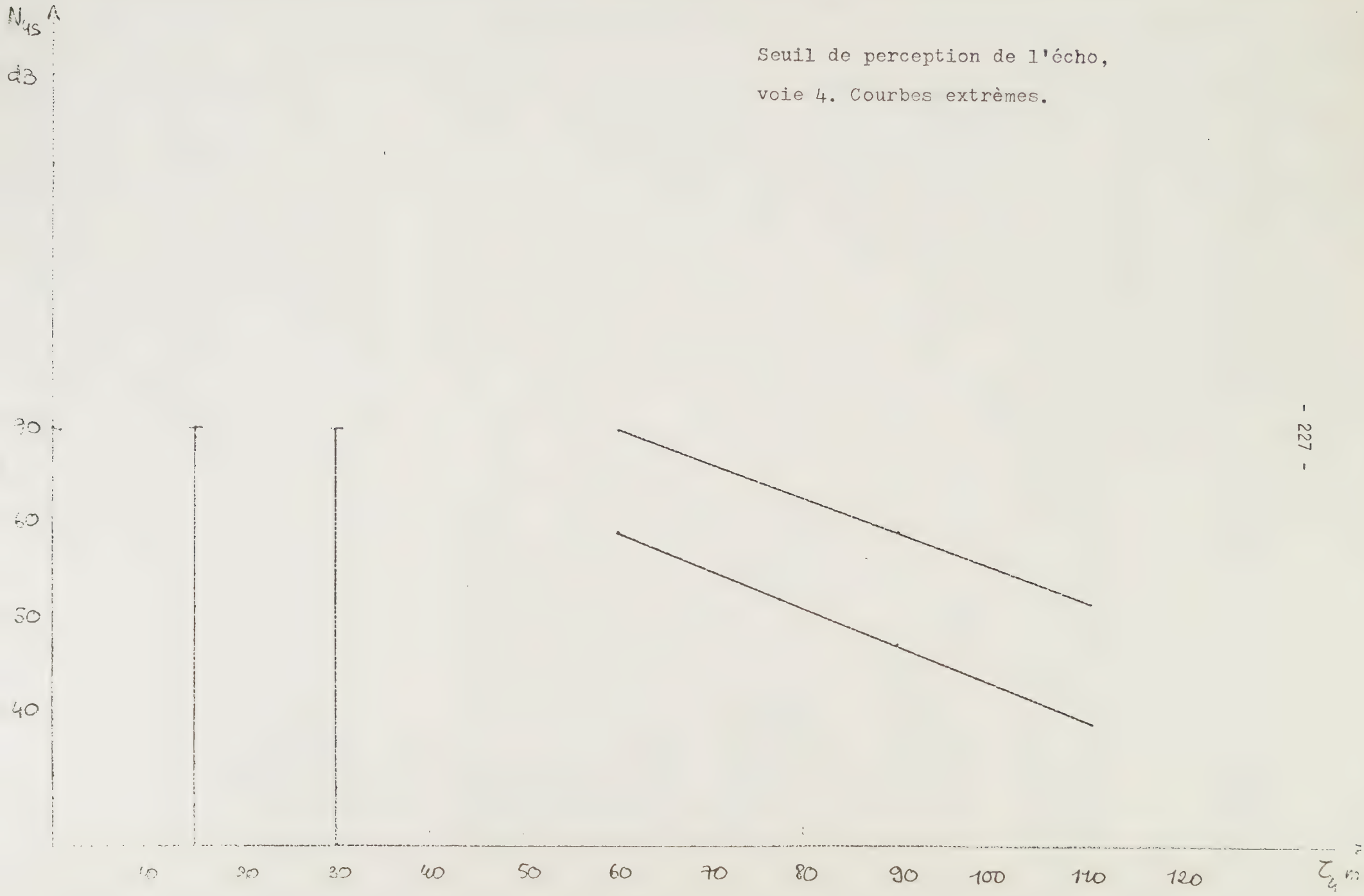
Seuil de perception de l'écho, voie 4.

$\theta_1 = -30°$

$\theta_3 = 60°$

$\theta_4 = 120°$





Seuil de perception de l'écho,

voie 4. Courbes extrèmes.

son direct qui, par son intensité, son retard et sa direction, détermine la courbe N_{4S} (τ_4). Examinons maintenant quelques situations particulières de θ_2 et θ_3.

α) Situation pseudo-symétrique.

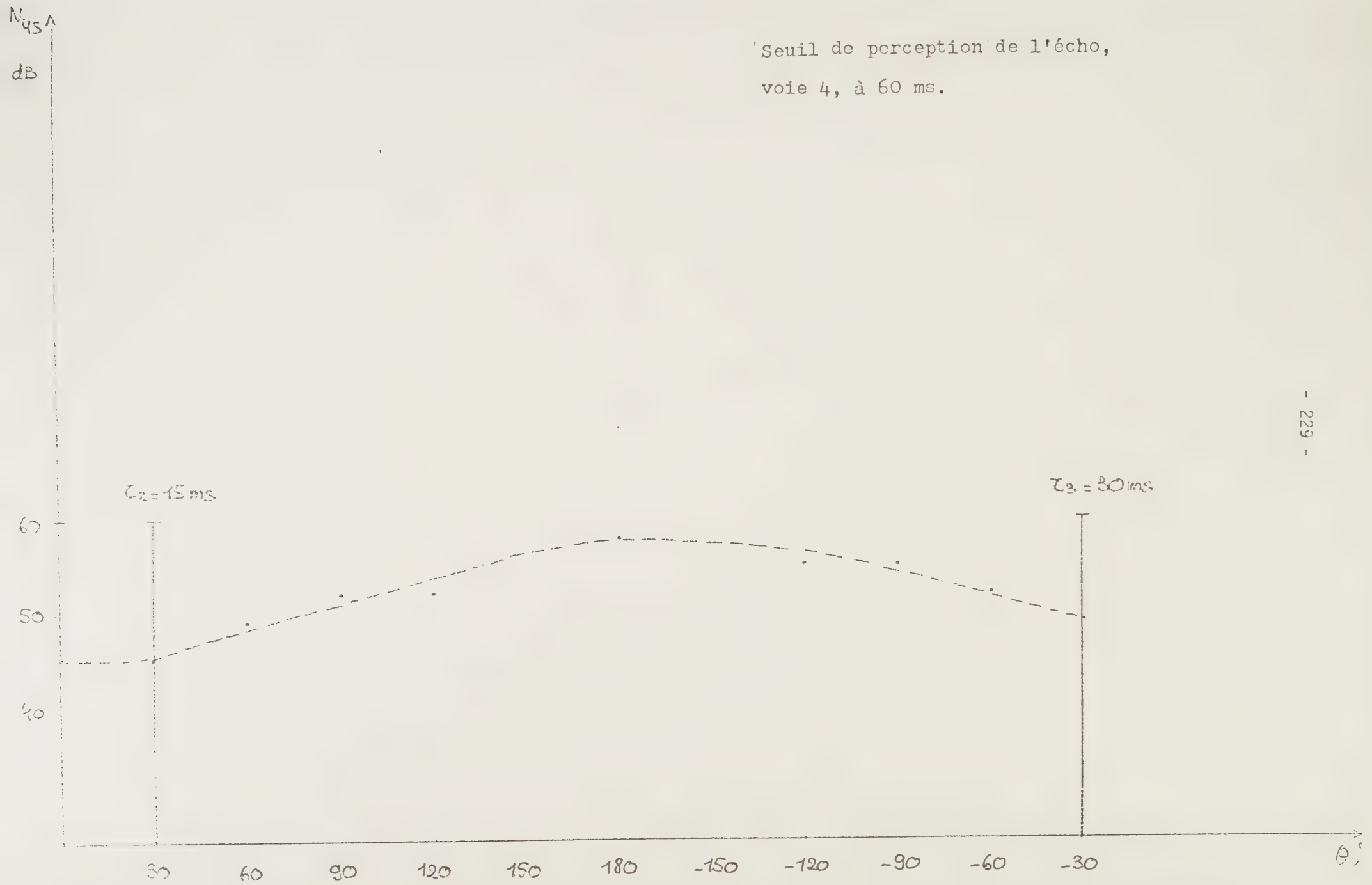
Si $\theta_1 = -\theta_2$, on a une situation géométrique analogue à celle envisagée plus haut lorsque la source primaire était tridirectionnelle, soit $\theta_{11}=0°$ et $\theta_{12} = -\theta_{13}$. Cette situation correspond aux seuils N_{2S} (θ_2) pour τ_2 = 30ms figurés sur les courbes des figures 5.19 et 5.20.
On peut alors comparer les nouvelles courbes des figures 5.37 à 5.40, qui correspondent à $N_{4S}(\theta_4)$ pour τ_4 = 60 ms. Dans le cas d'une source assez étroite (Fig. 5.37 et 5.38), on observe que le seuil passe par un maximum pour une direction θ_4 opposée au centre de la source primaire, et par un minimum pour $\theta_4 = 0$.
Cet effet est bien le même que celui relevé sur les courbes des Fig. 5.19 et 5.20; cependant, il apparaît que la présence de délais différents pour les voies 2 et 3, soit 15 ms pour la voie 2 et 30 ms pour la voie 3,entraîne une dissymétrie de la courbe N_{4S} (θ_4), plus élevée du côté de la réflexion la plus retardée, c'est-à-dire celle donnée par la voie 3.
Pour la situation $\theta_2 = 90°$, $\theta_3 = -90°$ (Fig. 5.39), l'effet de masquage direct pour la direction $\theta_4 = 0$, sensible sur la Fig. 5.20 disparaît, alors qu'un effet de masquage pour les directions $\theta_4 = 90°$ et $\theta_4 = -90°$ se confirme avec d'autant plus de netteté que la réflexion associée au son direct est plus retardée, c'est-à-dire pour $\theta_4 = -90°$. La figure 5.40 donne la courbe obtenue pour $\theta_2 = -0_3 = 120°$ et fait apparaître, au contraire de la Fig.5.20, un effet de masquage direct pour la direction $\theta = 180°$. La courbe N_{4S} (θ_4) accuse une dissymétrie, étant décalée du côté de θ_3.

β) Source primaire bilatérale dissymétrique.

Ce cas est représenté par les figures 5.41 et 5.42, et confirme les résultats obtenus ci-dessus. Cependant, il n'y a pas de différence très sensible avec le cas de la Fig. 5.21. On remarque le seuil particulièrement faible dans la région de la source primaire, avec un décalage vers la direction θ_2 ou θ_3 la plus éloignée de 0°. Cet effet est plus sensible que celui entraîné par les différences de délais entre les voies 2 et 3.

Fig. 7

'Seuil de perception' de l'écho,
voie 4, à 60 ms.



Seuil de perception de l'écho,
voie 4, à 60 ms.

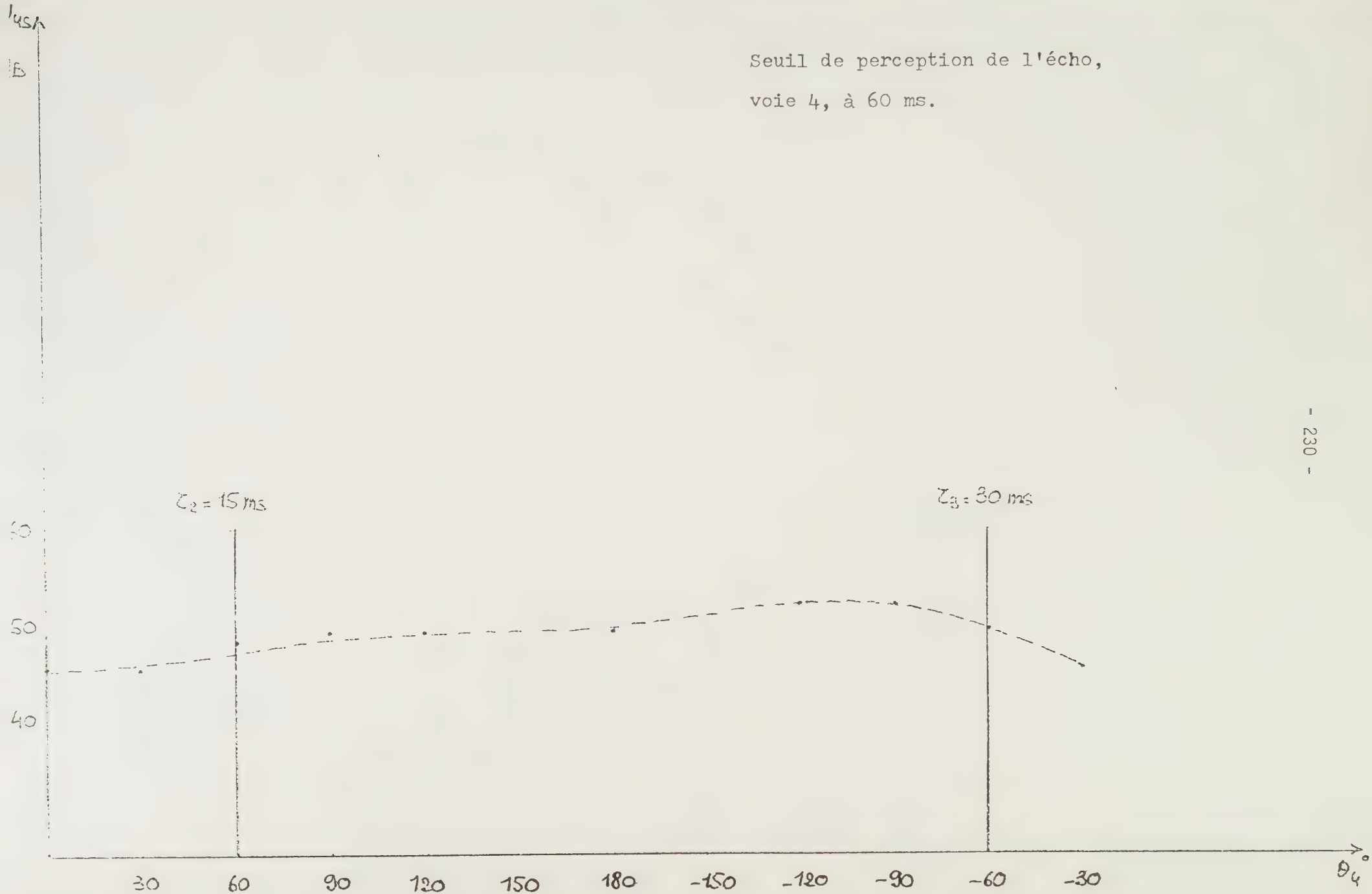


Fig 5.39

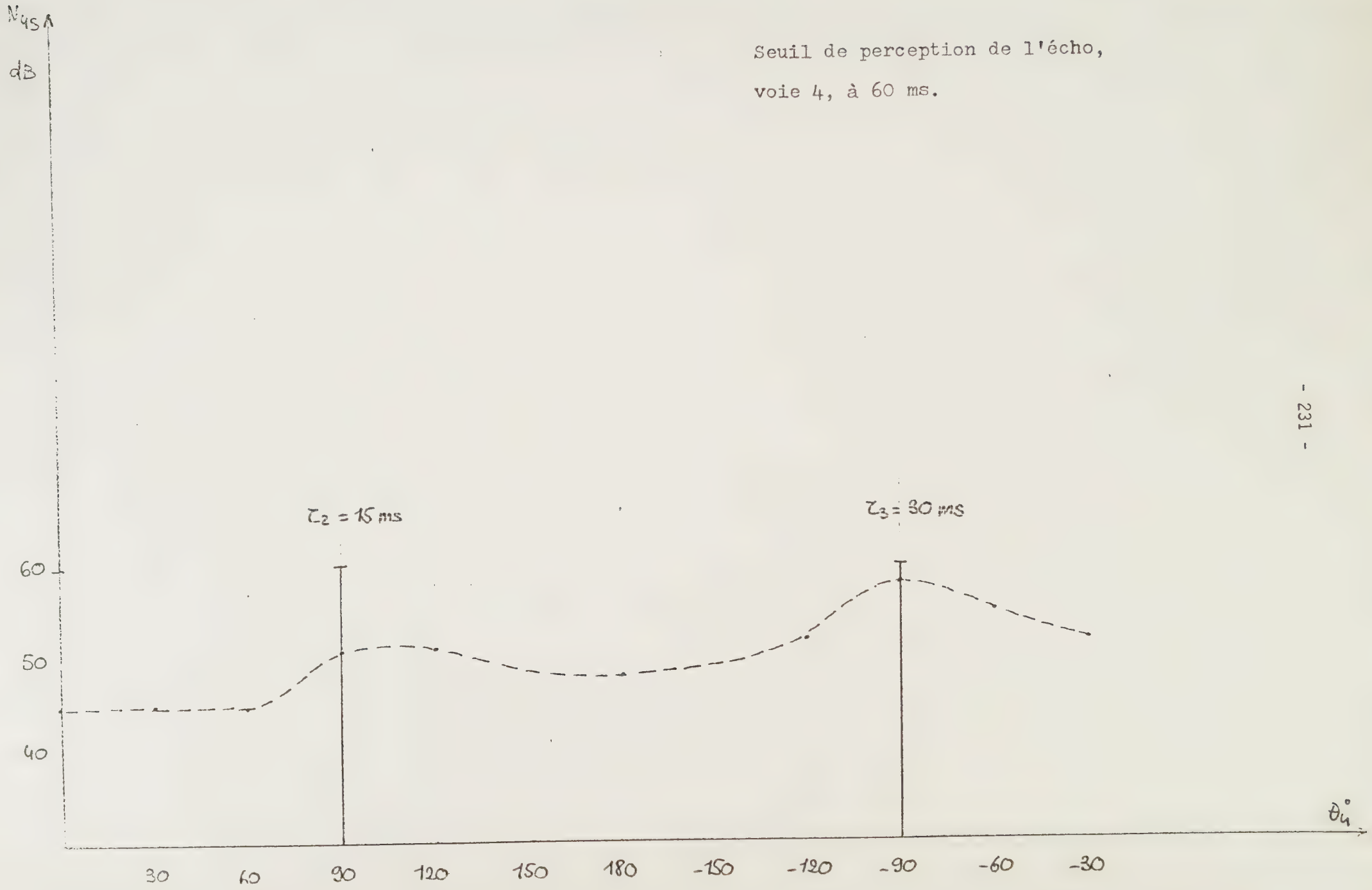
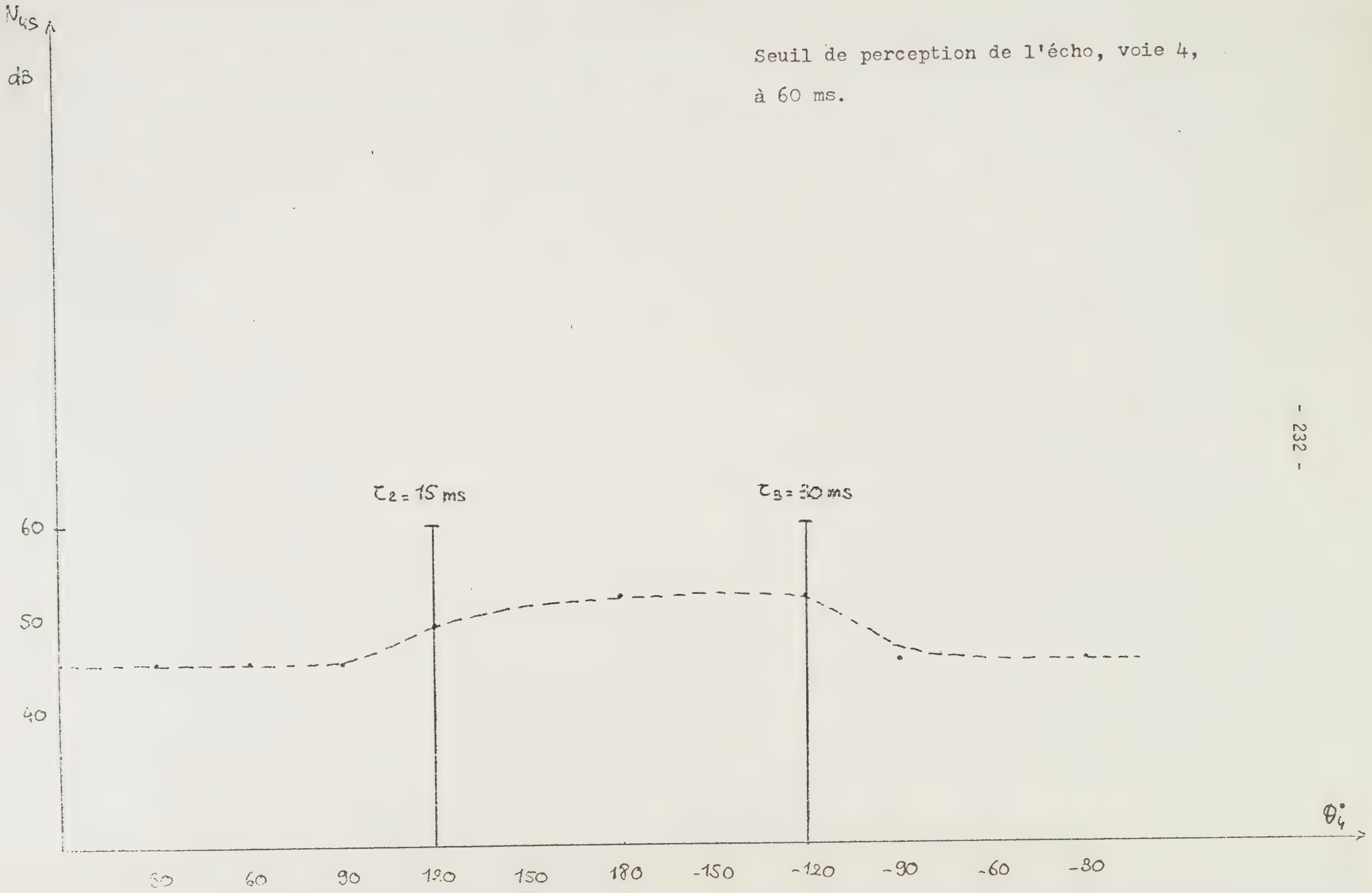
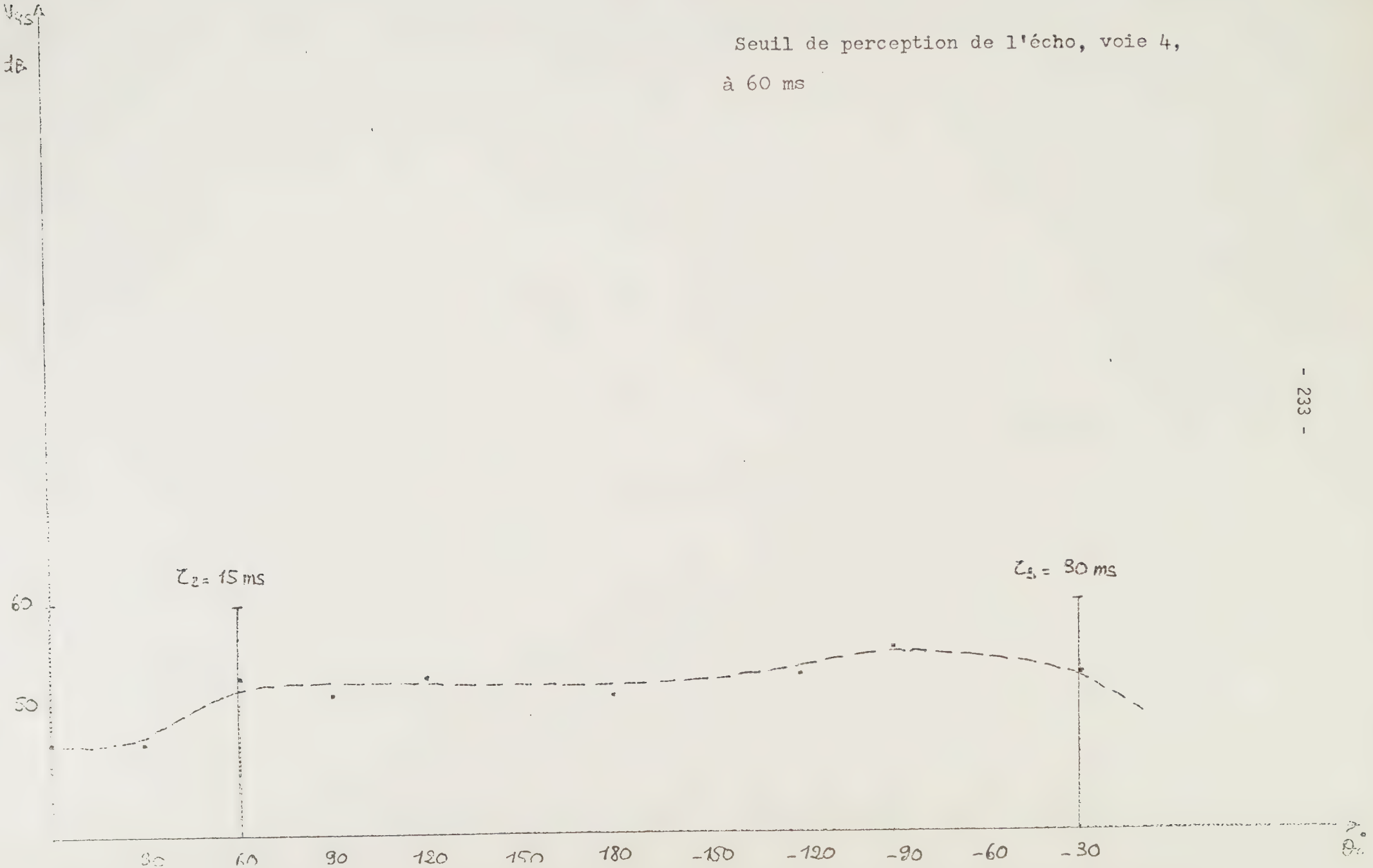

N_{4S}
dB
Seuil de perception de l'écho,
voie 4, à 60 ms.
τ_2 = 15 ms
τ_3 = 30 ms
60
50
40
30
60
90
120
150
180
-150
-120
-90
-60
-30
θ_4°

Fig 5.40

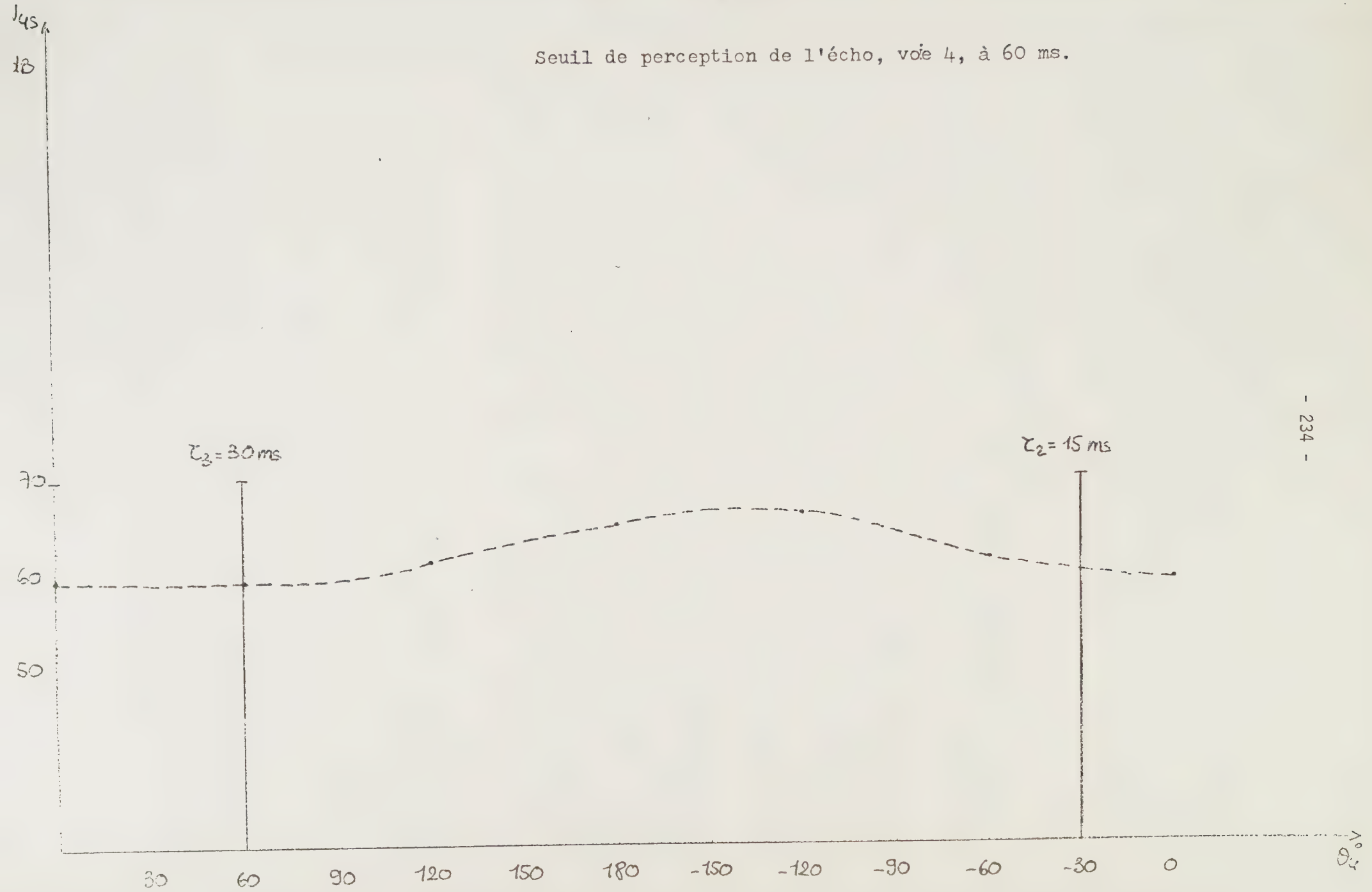

N_{4S}
dB
Seuil de perception de l'écho, voie 4,
à 60 ms.
$\tau_2 = 15$ ms
$\tau_3 = 30$ ms
60
50
40
30
60
90
120
150
180
-150
-120
-90
-60
-30
θ_4

Seuil de perception de l'écho, voie 4,

à 60 ms



Seuil de perception de l'écho, voie 4, à 60 ms.



γ) Source primaire monolatérale.

De la même manière, on peut comparer les résultats obtenus pour $\theta_2 = 30°$ et $\theta_3 = 60°$ (Fig. 5.43) avec ceux de la figure 5.23 ($\theta_{12} = 30°$, $\theta_{13} = 60°$). La présence des deux réflexions retardées semblent créer une zone de seuil minimal beaucoup plus large et faire disparaître la localisation très accusée du minimum.

- b) Source primaire avec dédoublements.-

Envisageons maintenant une source primaire constituée par le son direct de direction $\theta_1 = 0°$, deux réflexions de directions θ_{21} et θ_{22} à 15 ms et deux réflexions de directions θ_{31} et θ_{32} à 30 ms.

Si la situation de la source primaire est symétrique, avec par exemple $\theta_{21} = -\theta_{22} = 30°$ et $\theta_{31} = -\theta_{32} = 60°$, le seuil N_{4S} (θ_4) est donné par la figure 5.44. On remarque un maximum très accusé pour la direction $\theta_4 = 0$, provoqué par un effet de masquage direct intense.
En effet, les voies 2 et 3 créeent deux sources fantômes de direction $\theta = 0$. Le seuil N_{4S} est maximum pour $\theta_4 = 180°$.

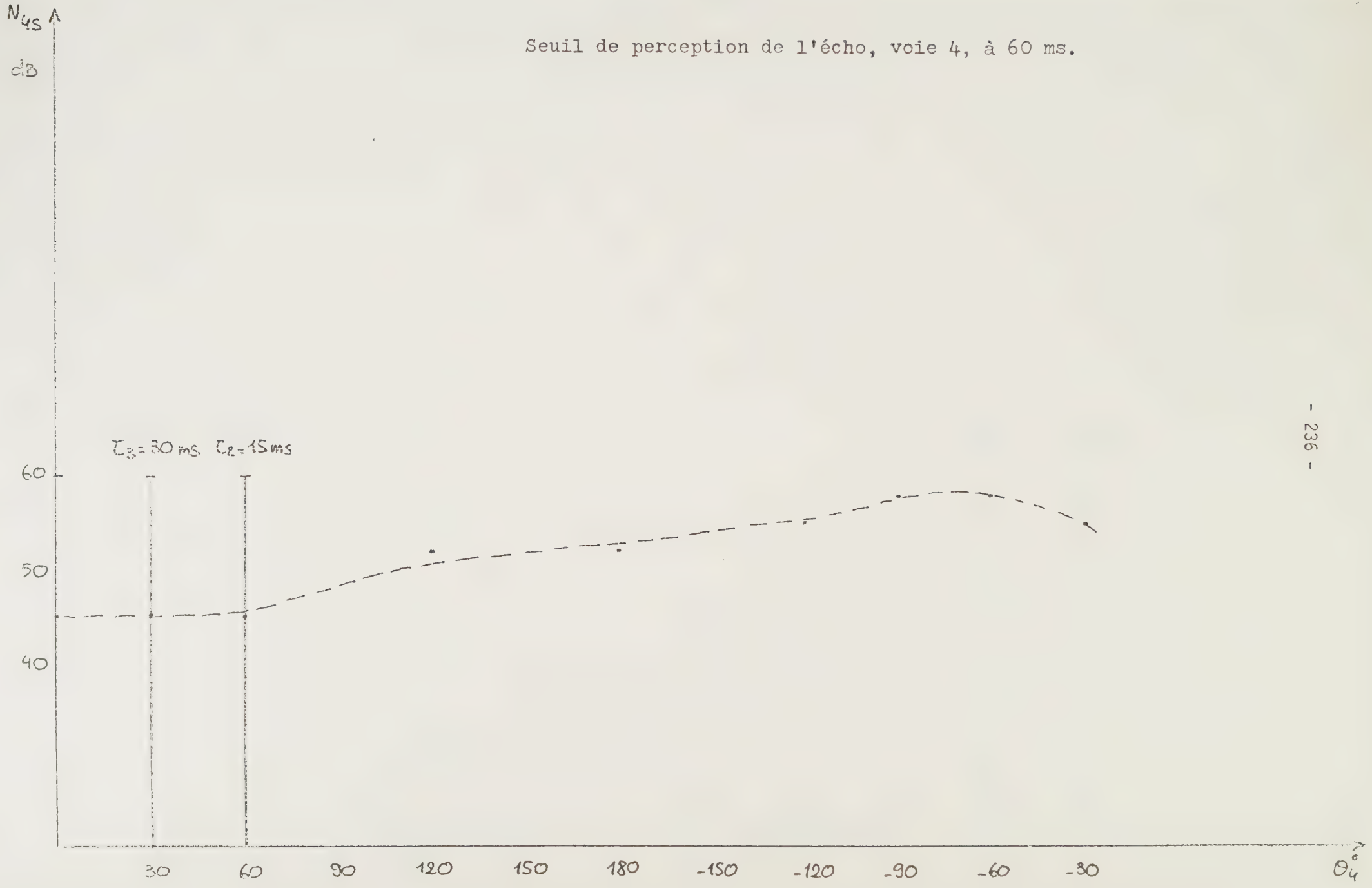
Si l'on envisage maintenant une source légèrement dissymétrique, par exemple avec $\theta_{21} = 60°$, $\theta_{22} = -30°$ et $\theta_{31} = 30°$, $\theta_{32} = -60°$, on constate que l'effet intense de masquage direct a disparu : le seuil est maximal lorsque la direction de l'écho est opposée à celle de la source primaire : il existe un léger décalage du minimum vers la source fantôme correspondant à la voie la moins retardée (Fig. 5.45). La figure 5.46 donne les courbes extrêmes obtenues pour cette situation de la source primaire.

5.6.- Utilisation des courbes de seuil de perception des échos pour la prévision de l'intelligibilité.-

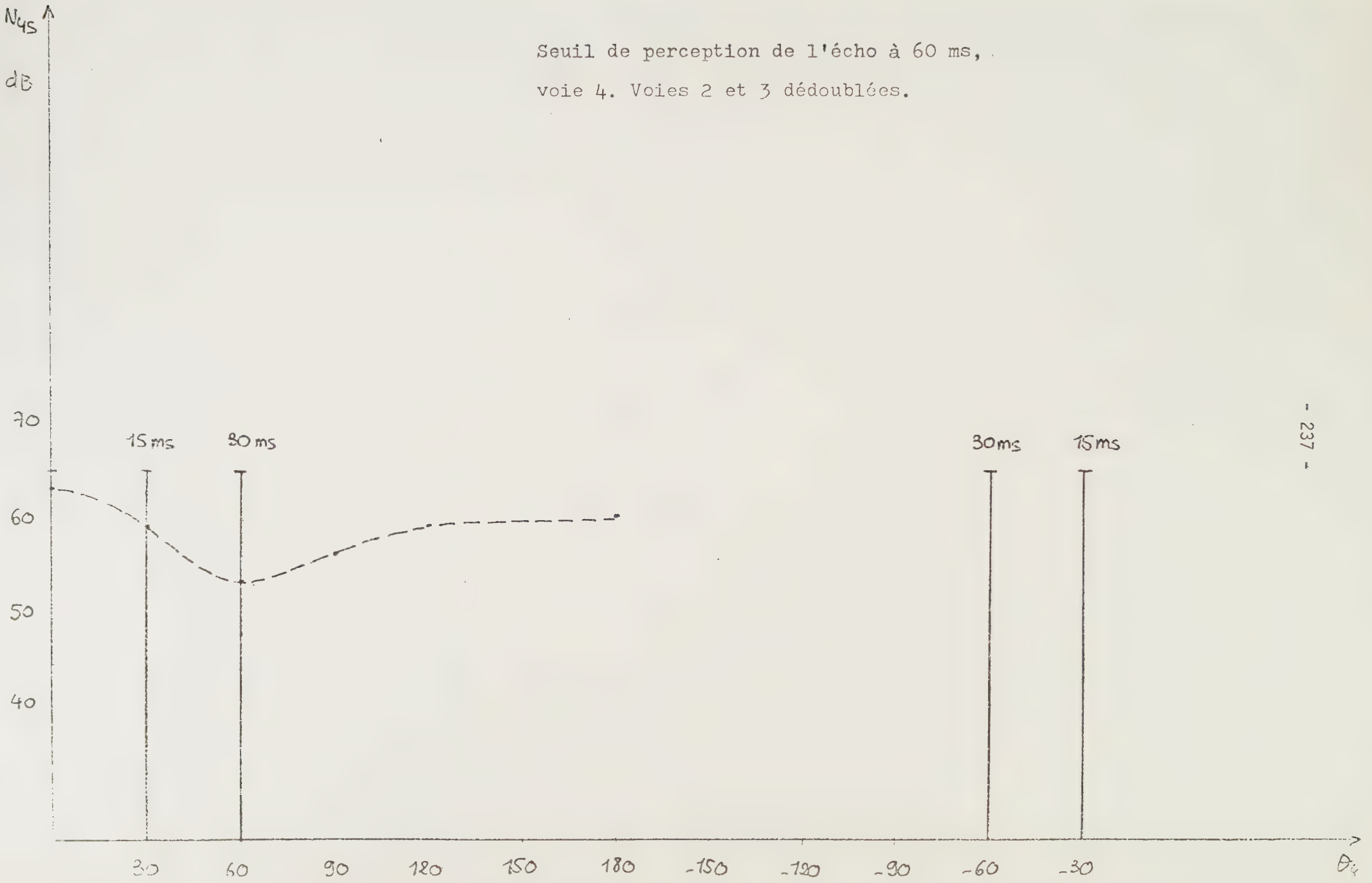
5.6.1.- Généralités.-

Les courbes expérimentales tracées ci-dessus ont pour but de remplacer les courbes "d'intégration auditive" utilisées par Niese par exemple et qui ne nous donnaient pas satisfaction. Etant donné un échogramme spatio-temporel, il faut évaluer quelle est l'énergie utile, l'énergie nuisible, ainsi que les directions respectives des deux énergies.

Seuil de perception de l'écho, voie 4, à 60 ms.

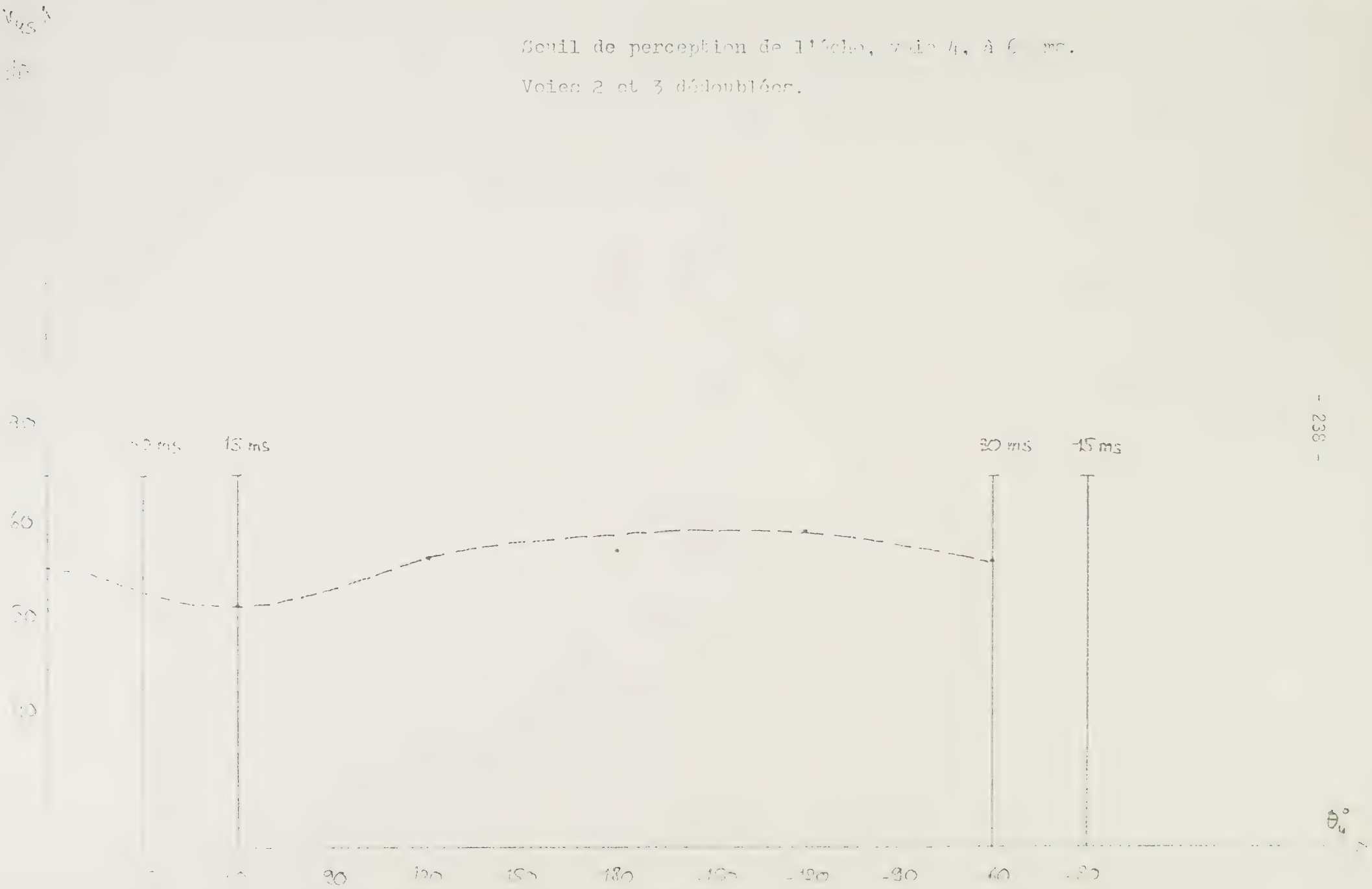


Seuil de perception de l'écho à 60 ms, voie 4. Voies 2 et 3 dédoublées.

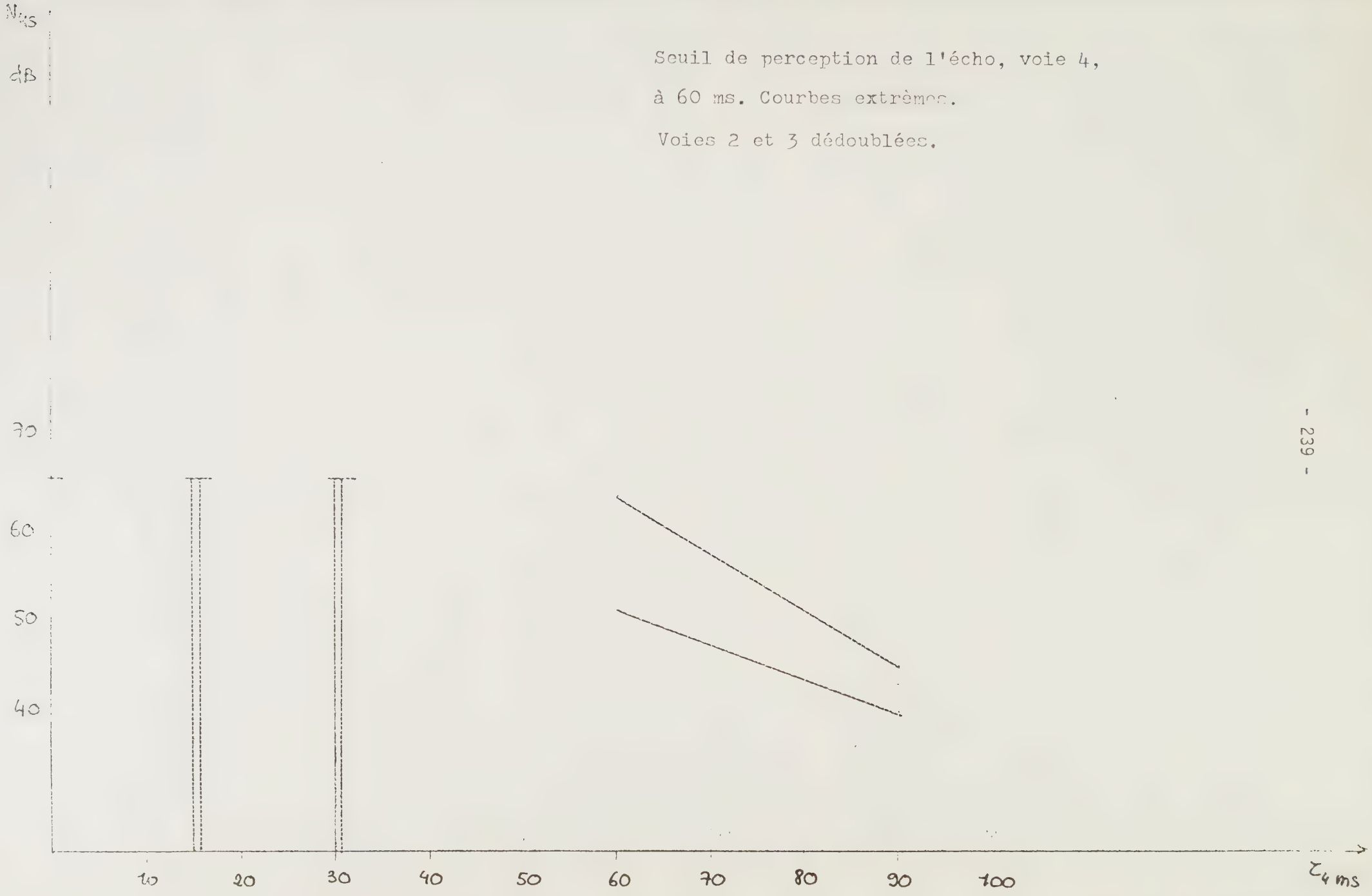


Seuil de perception de l'écho, voie 4, à 6 ms.

Voies 2 et 3 dédoublées.



Seuil de perception de l'écho, voie 4,
à 60 ms. Courbes extrêmes.
Voies 2 et 3 dédoublées.



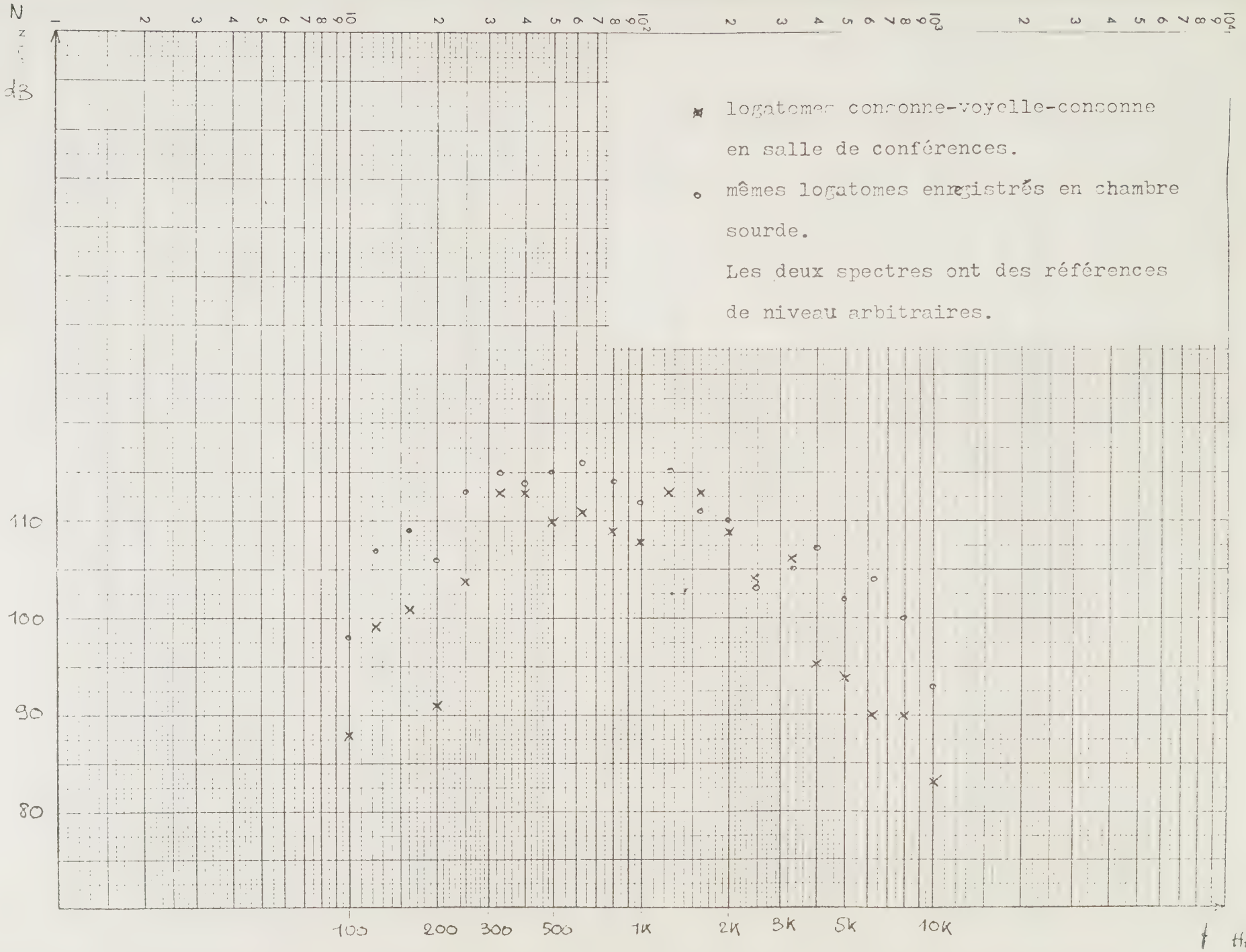
Nous procéderons alors de la façon suivante : le son direct a une intensité I_1 et une direction $\theta_1 = 0$: sa contribution à l'énergie utile est égale à I_1, alors que celle à l'énergie nuisible est nulle, soit $I_{1n} = 0$.
La première réflexion, d'intensité I_2 et de retard τ_2, aura des contributions $I_{2u} = I_{2S}(\tau_2)$ et $I_{2n} = I_2 - I_{2S}(\tau_2)$ si $I_2 > I_{2S}$ et $I_{2n} = 0$ si $I_2 \leqslant I_{2S}$.
Ainsi, pour la deuxième réflexion, la source primaire aura pour intensité $I_1 + I_{2u}$, comme directions O_1 et O_2. On en déduit alors une courbe $I_{3S}(\tau_3)$ qui permet de connaître $I_{3u} = I_{3S}(\tau_3)$ et I_{3n}, suivant le même schéma que celui utilisé pour I_2.
L'énergie nuisible éventuelle aura pour intensité $I_{2n} + I_{3n}$, et pour directions θ_2 et θ_3. Energie utile et énergie nuisible ont la particularité d'appartenir l'une et l'autre au même type de signal acoustique : il s'agit de parole, dont le spectre peut être éventuellement altéré par des réflexions de coefficient variable avec la fréquence.

Pour simplifier les choses, nous supposerons que le spectre de la parole réfléchie est identique à celui de la parole émise : naturellement, c'est ce qui se passe dans la manipulation du champ sonore synthétique que nous utilisons, cependant, il est possible de modifier le spectre du signal donné par telle voie pour simuler par exemple l'atténuation des hautes fréquences données par les réflexions successives sur les parois. A titre d'exemple, on a (Fig. 5.47) le spectre d'un échantillon de parole enregistré en chambre sourde, et celui obtenu après passage de cet échantillon dans une salle dont la réverbération est de l'ordre de 0,5 s à 1000 Hz (salle de conférences des laboratoires du CNRS à Marseille).

L'énergie nuisible est donc un "bruit", c'est-à-dire un signal acoustique indésirable, de spectre analogue à celui de la parole. Il nous est apparu nécessaire d'effectuer une série de mesures d'intelligibilité pour étudier l'influence de la direction d'un bruit sur la masquage de la parole.

5.6.2.- Mesures d'intelligibilité en présence de bruit.-

Nous avons indiqué ailleurs quels sont les éléments phonétiques utilisés pour nos mesures d'intelligibilité : il s'agit de logatomes du type Consonne-Voyelle-Consonne, insérés dans de courtes phrases introductives et conclusives. Ils sont au nombre de 1000, soit 20 listes


N
dB
1
2
3
4
5
6
7
8
9
10
10^2
10^3
10^4
110
100
90
80
100
200
300
500
1K
2K
3K
5K
10K
f Hz
x logatomes consonne-voyelle-consonne en salle de conférences.
o mêmes logatomes enregistrés en chambre sourde.
Les deux spectres ont des références de niveau arbitraires.

de 50 logatomes. Chaque liste est précédée par un son de référence à 1000 Hz dont le niveau est situé à 6 dB au-dessous des maxima du signal de la parole. Chaque mesure d'intelligibilité est effectuée sur un ensemble de 200 logatomes.

a) - Cas d'une source primaire monodirectionnelle.-

On envisage la voie 1 de niveau égal à 40 dB, et une source qui donne un bruit de spectre analogue à celui de la parole, de direction θ_B variable. La figure 5.48 donne les variations de l'intelligibilité en fonction de θ_B pour N_B = 40 dB et N_B = 50 dB.
Il apparaît nettement que c'est pour $\theta_B = 0$ que l'intelligibilité est la plus faible. Au contraire, elle passe par un maximum pour $\theta_B \simeq 120°$: pour $\theta_B = 180°$, elle passe par un minimum moins accusé que celui obtenu pour $\theta_B = 0$. L'écart dans les variations de l'intelligibilité est d'autant plus faible que le niveau du bruit de masque est moins élevé.

b) - Cas d'une source primaire bidirectionnelle.-

Si l'on envisage maintenant une source primaire constituée par deux voies en parallèle, de directions $\theta_{11} = 0$ et $\theta_{12} = 30°$, et dont l'intensité totale est de 43 dB, on obtient les courbes de la figure 5.49 en utilisant un bruit de masque de 43 ou 53 dB, de direction θ_B variable. Les courbes obtenues ressemblent assez à celles de la figure 5.48 : on observe cependant une nette augmentation de l'intelligibilité, en particulier pour la courbe relative à N_B = 53 dB. Pour les directions $\theta_{11} = 0°$ et $\theta_{12} = 60°$, la courbe de la figure 5.50 se signale à nouveau par une augmentation de l'intelligibilité pour $\theta_B = 0°$. Le maximum correspond à une direction $\theta_B = \theta_{12} - 180°$. La figure 5.51 donne les résultats obtenus lorsque la source primaire a les directions $\theta_{11} = 0°$ et $\theta_{12} = 120°$.

On constate que, d'une façon générale, la source de bruit est particulièrement perturbatrice si sa direction est la même que celle de la source primaire. Ce résultat recoupe ceux de Januska et Spronck [5.10], [Januska] [5.13] et Lehmann et Wettschureck [5.9]. On est par conséquent en présence d'une situation parallèle à celle du seuil de perception d'un écho : l'effet perturbateur, dû à une réflexion ou à un bruit est d'autant plus sensible que sa direction est proche de celle de la source primaire, et en particulier du son direct. Cet effet est dû à l'audition binaurale.

Intelligibilité de la parole
de source monodirectionnelle.

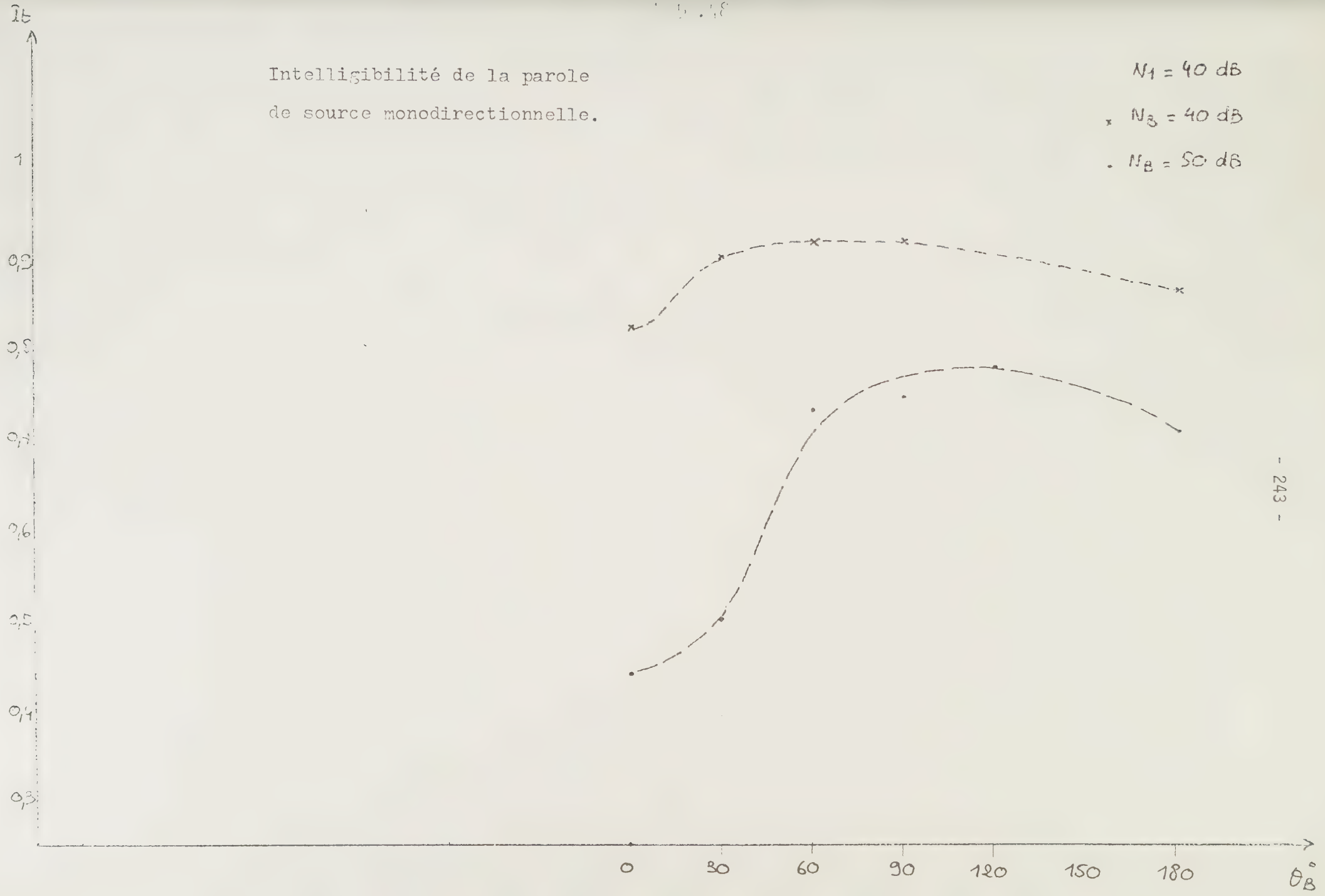
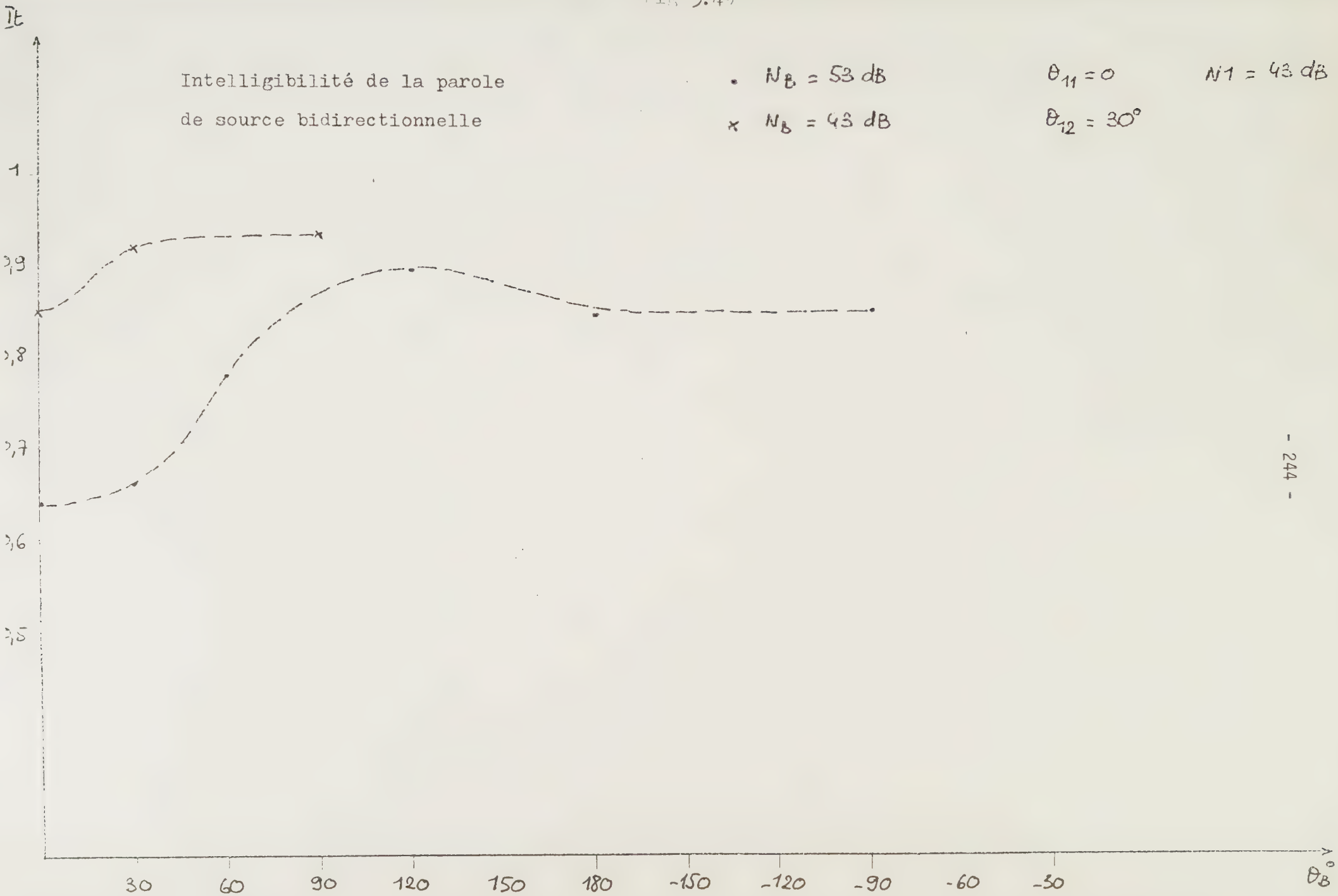


Fig 5.49


It
Intelligibilité de la parole
de source bidirectionnelle
• N_B = 53 dB
× N_B = 43 dB
θ_{11} = 0
θ_{12} = 30°
N1 = 43 dB
1
,9
,8
,7
,6
,5
30
60
90
120
150
180
-150
-120
-90
-60
-30
θ_B°

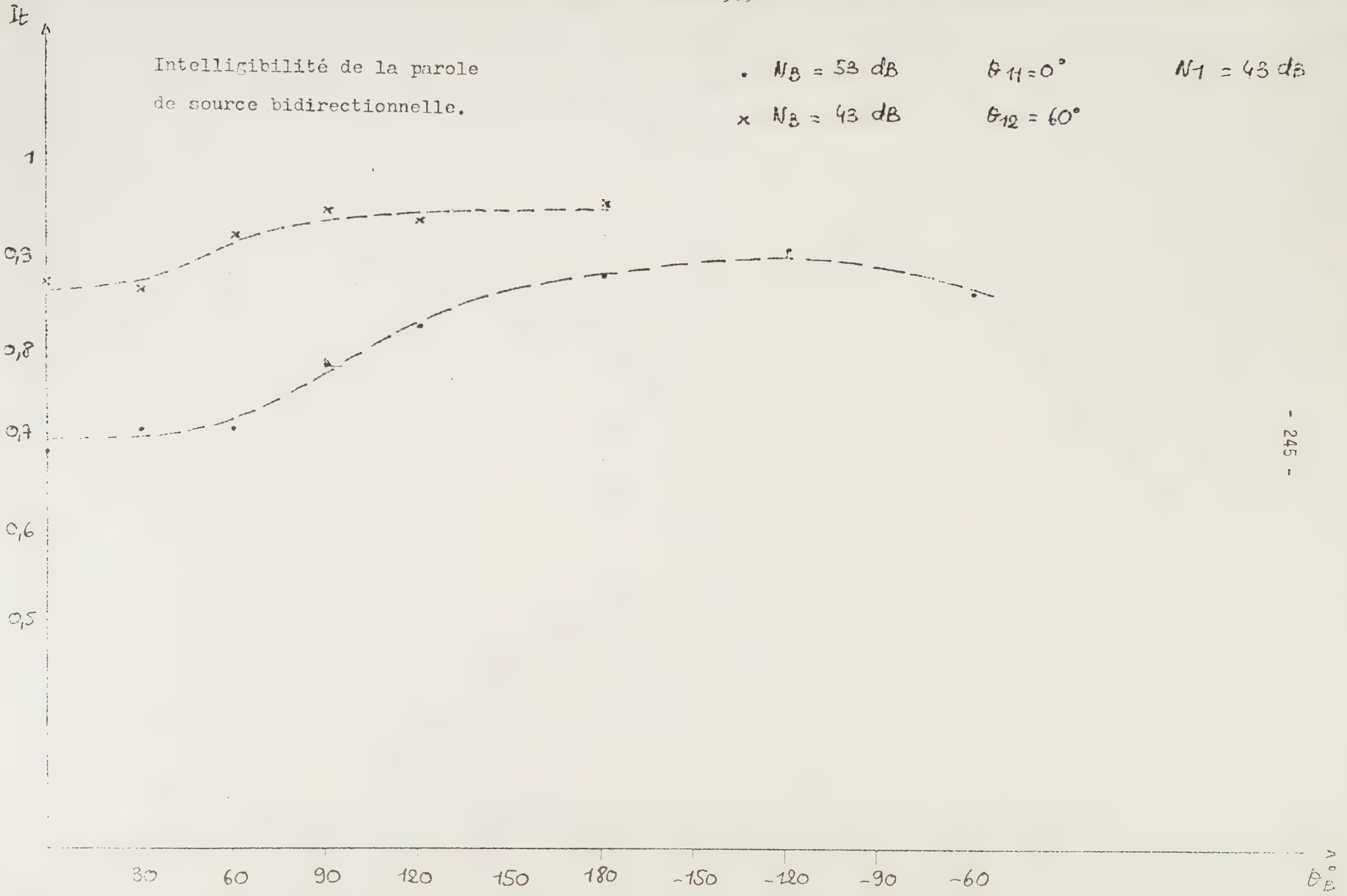
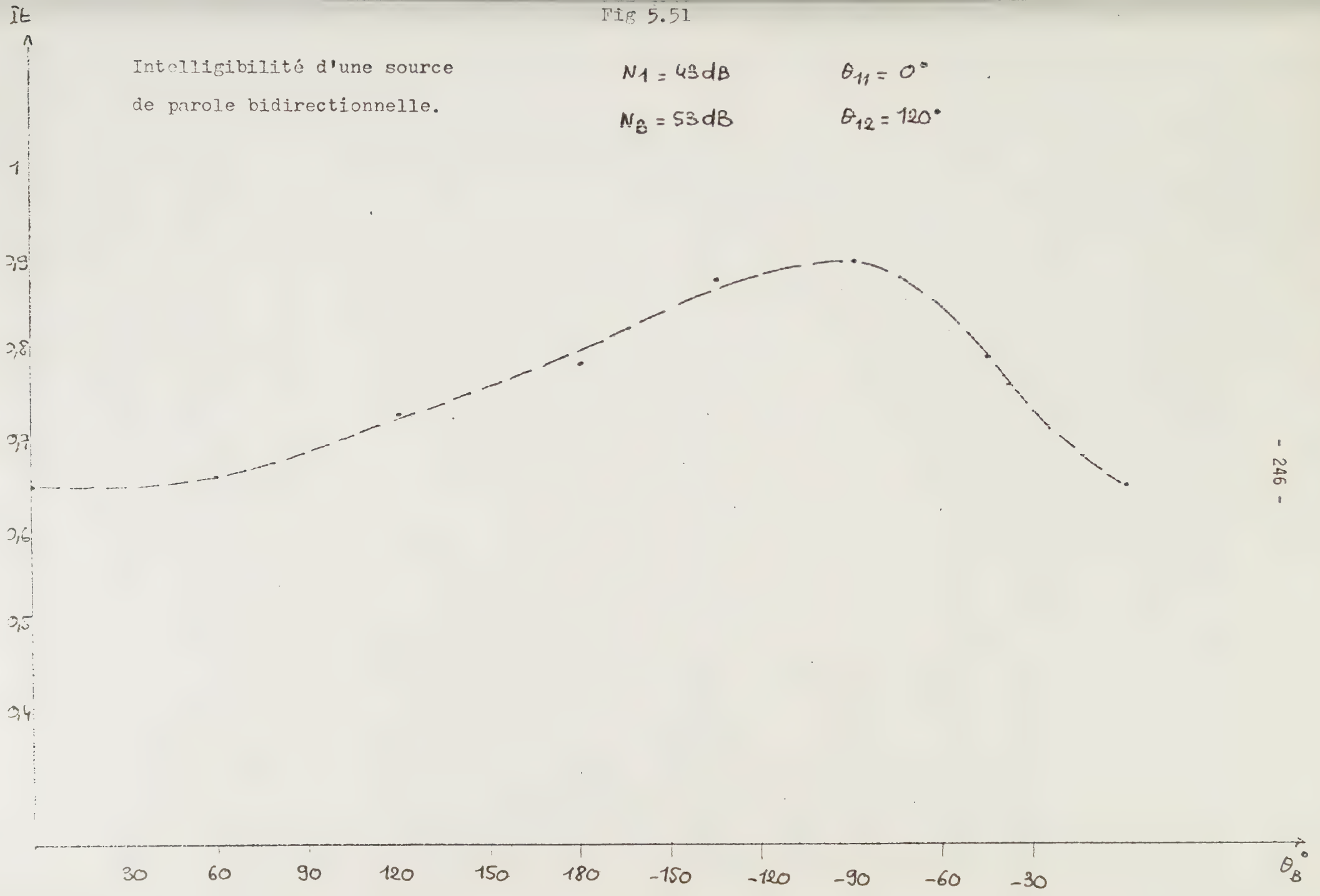

It
Intelligibilité de la parole
de source bidirectionnelle.
• N_B = 53 dB
× N_B = 43 dB
θ_{11} = 0°
θ_{12} = 60°
N_1 = 43 dB
1
0,9
0,8
0,7
0,6
0,5
30
60
90
120
150
180
-150
-120
-90
-60
θ_B

Fig 5.51

Intelligibilité d'une source de parole bidirectionnelle.

$N_1 = 43dB$ $\theta_{11} = 0°$

$N_8 = 53dB$ $\theta_{12} = 120°$



D'autre part, les variations It (θ_B) sont plus faibles si la source de parole est large : nous avons d'ailleurs remarqué que ces variations sont encore moindres si la source de bruit est elle aussi élargie, c'est-à-dire si par exemple elle est bidirectionnelle, la direction θ_{B1} restant fixe.

c) - Intelligibilité en présence d'un bruit diffus.-

Les différences obtenues pour le coefficient d'intelligibilité suivant les valeurs de l'angle θ_B nous ont amené à tracer des courbes "moyennes", dans une condition intermédiaire entre la plus favorable (valeur maximale de l'intelligibilité) et la plus défavorable (valeur minimale de l'intelligibilité). En fait, la source de bruit a une direction centrée sur $\theta = 0$, car les mesures d'intelligibilité en présence d'échos dont nous parlerons ultérieurement ont été effectuées avec un bruit de fond de direction frontale. La figure 5.1 présente ce réseau de courbes tracées dans les conditions suivantes : la source primaire est monodirectionnelle, avec $\theta_1 = 0$: la source de bruit est tridirectionnelle, avec $\theta_{B1} = 0$, $\theta_{B2} = 60°$, $\theta_{B3} = -60°$. Le niveau de bruit est pris comme paramètre, alors que le niveau de la source primaire N_1 est porté en abscisse. Chaque point correspond à une mesure effectuée à l'aide de 400 logatomes, soit 8 listes de 50 logatomes, ce qui demande 50 minutes environ. Nous verrons plus loin comment nous avons été amenés à utiliser de façon systématique ce réseau de courbes "moyennes" pour effectuer des prévisions d'intelligibilité dans le champ sonore synthétique. En tout cas, il faut remarquer que ces courbes ont une croissance assez rapide lorsque l'intelligibilité est inférieure à 80%, et nettement moins rapide au-delà. Par conséquent, si l'on veut avoir des différences sensibles entre les diverses situations acoustiques, c'est-à-dire entre les différents niveaux de l'énergie utile N_u et de l'énergie nuisible N_n, qui puissent entraîner des mesures d'intelligibilité assez différentes, on a tout intérêt à travailler avec des niveaux assez faibles.

5.6.3.- Cas de l'écho donné par la voie 1.-

Avant de traiter de ce cas très simple, il faut noter que les calculs portent sur les intensités des différentes voies. Si I_1 est l'intensité de la voie 1, son niveau N_1 est alors :

$$N_1 = 10 \log \ (I_1 / 10^{-16})$$

si le seuil auditif a pour intensité $I = 10^{-16}$ W/cm2, soit $p = 2.10^{-4}$ baries
Il vient alors :

$$I_1 = 10^{-16} \; e^{0,23\, N_1}$$

d'autre part, nous avons toujours supposé l'existence d'un bruit de fond. Si, dans la chambre sourde, le bruit de fond est très réduit, réduit en fait à celui de la bande magnétique, tel n'est pas le cas dans une salle réelle. C'est pourquoi nous avons toujours utilisé un bruit de fond artificiel, dont, pour des raisons de simplicité, le spectre a été pris identique à celui de la parole. Soit N_B le niveau de bruit.

Envisageons la disposition suivante :

$N_1 = N_2 = N_B = 30$ dB

$\theta_2 = 0 \quad \tau_2 = 30$ ms

$\theta_B = 0$

D'après la courbe de seuil tracée plus haut (Fig. 5.10) le niveau N_{2S} sera égal à 24 dB

On a alors le bilan suivant :

$$I_u = I_1 + I_{2u} = 10^{-16} \left(e^{0,23.30} + e^{0,23.24} \right)$$

$$= 10^{-16} \left(1000 + 245 \right)$$

On en déduit $N_u = 31$ dB. L'intensité nuisible est celle du bruit de fond, augmentée de l'intensité nuisible de l'écho :

$$I_n = I_B + I_2 - I_{2u} = 10^{-16} \left(e^{0,23.30} + e^{0,23.30} - e^{0,23.24} \right)$$

$$N_n = 32,5 \text{ dB}$$

En se reportant aux courbes de la Fig. 5.1, on obtient alors un coefficient d'intelligibilité de 79%. Une mesure directe, dans les mêmes conditions acoustiques, a donné un coefficient de 80%. On peut faire les remarques suivantes :

- Un son dont le niveau est inférieur de 6 dB à un autre le renforce à peine, puisque le décibel correspond au seuil de discrimination entre deux intensités : on conçoit ainsi la différence qui peut exister entre une ré-

flexion non intégrée au son direct - ou intégrée très partiellement - et entre une réflexion complétement intégrée, c'est-à-dire entre les deux situations :

$$N_u = 31 \text{ dB}, N_n = 32{,}5 \text{ dB} \quad , \text{ soit } It = 79\% \text{ et}$$

$$N_u = 33 \text{ dB}, N_n = 30 \text{ dB} \quad , \text{ soit } It = 87\%$$

- D'une façon rigoureuse, il ne faudrait pas utiliser le réseau de courbes de la Fig. 5.1, mais un réseau particulier pour chaque situation de la source primaire et de la source de bruit : en particulier, dans le cas précédent, il aurait fallu avoir un signal de parole d'intensité $I_u = I_1 + I_{2u}$ de direction $\theta_1 = 0$, et un signal de bruit (nuisible) d'intensité $I_n = I_B + I_2 - I_{2u}$, et de direction $\theta_B = 0$. Bien entendu, le résultat ne serait pas rigoureusement le même; cependant, pour des raisons d'ordre pratique, nous avons été amenés à simplifier les choses et à utiliser un seul réseau de courbes. Mais nous pensons que la simplification ainsi apportée ne provoque qu'une erreur assez faible sur l'estimation de l'intelligibilité: comparons par exemple les situations très différentes des figures 5.48 et 5.50, soit :

$\theta_B = 0$, $\theta_1 = 0$ d'une part, et

$\theta_B = 0$, $\theta_{11} = 0$, $\theta_{12} = 60°$ d'autre part, avec

$N_1 = N_B = 43$ dB dans les deux cas : l'intelligibilité est de 82,5% pour la première situation, de 86% pour la seconde. Cette différence est relativement faible par rapport à celle provoquée par différents niveaux du seuil de perception.

Considérons maintenant la situation suivante :

$N_1 = N_2 = N_B = 30$ dB

$\theta_2 = 0$ $\quad \tau_2 = 150$ ms

$\theta_B = 0$

Pour cette valeur de τ_2, le niveau N_{2S} est au moins à 10 dB au-dessous

de N_1 . Dans ce cas, on obtient :

$I_u = I_1$ soit $N_u = 30$ dB

$I_n = I_2 + I_B$ soit $N_n = 33$ dB

En se reportant au réseau des courbes de la fig. 5.1, on obtient une valeur It_c calculée égale à 70%. La valeur de l'intelligibilité mesurée dans la même situation est de 68%.

Si l'on considère maintenant la situation :

$N_1 = N_2 = N_B = 30$ dB

$\theta_2 = 30°$, $\tau_2 = 30$ ms

$\theta_B = 0°$

on a vu précédemment que cette réflexion pouvait être considérée comme parfaitement intégrée au son direct. Dans ce cas, $N_u = 33$ dB et $N_n = 30$ dB. L'intelligibilité"calculée" est alors $It_c = 87\%$. Celle mesurée dans les conditions expérimentales, It_e est de 89%.
Le tableau 1 résume les résultats obtenus pour diverses valeurs de θ_2, τ_2 et N_2 . Les valeurs des intelligibilités expérimentales et calculées sont notées It_e et It_c. L'intelligibilité calculée est déduite des valeurs de N_u et N_n et du réseau de courbes de la figure 5.1.

D'une façon générale, il apparaît une bonne concordance entre les valeurs expérimentales et calculées. Il faut remarquer d'ailleurs que les valeurs It_e risquent d'être d'autant moins précises que la valeur de l'intelligibilité est plus faible. Ces mesures montrent que le seuil de perception de l'écho est bien lié à une nuisance éventuelle, et qu'il détermine l'évaluation correcte de l'énergie utile et de l'énergie nuisible.

5.6.4.- Cas de plusieurs échos.-

Le processus est toujours le même : le son direct, d'intensité I_1 amène une contribution à l'énergie utile égale à I_1 : la première réflexion, de direction θ_2, φ_2 d'intensité I_2 et de retard τ_2 apporte une contribution $I_{2u} = I_{2S}$ à l'énergie utile, avec $I_{2u} \leq I_2$. La contribution

Tableau 1

$N_1 = N_2 = N_B$

θ_2°	τ_{2ms}	N_1 dB	$N_{2s} - N_1$ dB	It_c %	It_e %
0	30	30	-6	79	80
0	150	30	-10	70	68
30	30	30	0	87	89
30	150	30	-10	72	75
30	50	30	-5	80	84
30	30	40	0	91	92
30	50	40	-5	88	85
30	80	40	-13	82	82
0	30	40	-6	85	85
0	50	40	-10	83	86
0	80	40	-13	82	83

à l'énergie nuisible est alors $I_{2n} = I_2 - I_{2u}$.

La source primaire d'intensité $I_1 + I_{2u}$, de direction θ_1, φ_1 et θ_2, φ_2 détermine alors un seuil I_{3S} , pour la deuxième réflexion de retard τ_3 et de directions θ_3, φ_3. L'énergie utile totale sera alors :

$I_u = I_1 + I_{2u} + I_{3u}$ et l'énergie nuisible

$I_n = I_{2n} + I_{3n}$ avec $I_{2n} = I_2 - I_{2S}$

$I_{3n} = I_3 - I_{3S}$

La valeur I_{2S} est connue à l'aide des courbes des figures 5.10 et 5.12.

Suivant la valeur de I_{2S}, soit très proche de I_1 (au cas où I_2 est pratiquement égal à I_1), soit au contraire beaucoup plus faible (au moins à 10 dB au-dessous de I_1), on pourra utiliser les courbes des figures 5.26 à 5.32 ou les courbes relatives au seuil de perception du premier écho. Si l'intensité I_2 est inférieure de quelques décibels à I_1, c'est une courbe moyenne entre celles relatives aux échos donnés par les voies 2 ou 3 qui pourront nous renseigner sur la situation de I_{3S} , par exemple la courbe de la figure 5.33.

Le tableau 2 donne quelques exemples de résultats obtenus en utilisant deux réflexions.
Comme dans le cas du premier écho, on peut estimer que la concordance entre l'intelligibilité mesurée et la prévision est très satisfaisante.

5.7.- Conclusion.-

Cet ensemble de mesures ne permet pas de dire qu'il existe une "formule" permettant de prévoir l'intelligibilité dans une salle : tel n'était pas notre propos, telle n'était pas notre prétention. Nous avons voulu mettre en évidence le rôle de la direction des échos qui composent un échogramme, et d'autre part établir une relation certaine entre le seuil de perception des échos et l'intelligibilité .

Quelle conclusion en tirer dans la pratique?
Si l'on a affaire à une salle de petites dimensions, le son direct est sui-

Tableau 2

$N_1 = N_2 = N_3 = N_B = 40$ dB
$\tau_2 = 30$ ms
$\tau_3 = 60$ ms

θ_2 °	θ_3 °	$N_{3S} - N_1$ dB	It_c %	It_e %
30	0	-5	90	89
30	60	-5	90	90
30	120	0	93	93
60	0	-15	88	92,5
60	30	-15	88	88
60	180	-4	91	91
60	-90	0	93	93
90	-90	-3	91	90
90	0	-10	88	87,5

vi de réflexions assez proches : la première intervient moins de 15 ms après le son direct, comme dans l'exemple de la figure 5.52 : dans ce cas, relatif à une salle dont toutes les parois, sauf le sol, sont très réfléchissantes, on peut estimer que l'énergie provient de tous les côtés. Quant à celle des réflexions retardées, amenant une énergie nuisible, elle provient également de directions très diverses, si bien que le rôle propre de la direction des réflexions devient faible. Nous avons fait des estimations d'intelligibilité, en utilisant la netteté au sens de Thiele : ces estimations se sont avérées très satisfaisantes dans le cas de cette petite salle. On peut s'étonner que ce "découpage" grossier qui sépare l'énergie utile et l'énergie nuisible par une frontière verticale à 50 ms puisse donner d'aussi bons résultats. En fait, il suffit de la comparer aux courbes de la Fig. 5.53 pour comprendre que la situation de la frontière n'est pas invraisemblable : elle est située entre la courbe de perception de l'écho donné par la voie 2, et celle relative à la voie 3, lorsqu'on a associé au don direct une première réflexion à 30 ms et de même intensité.

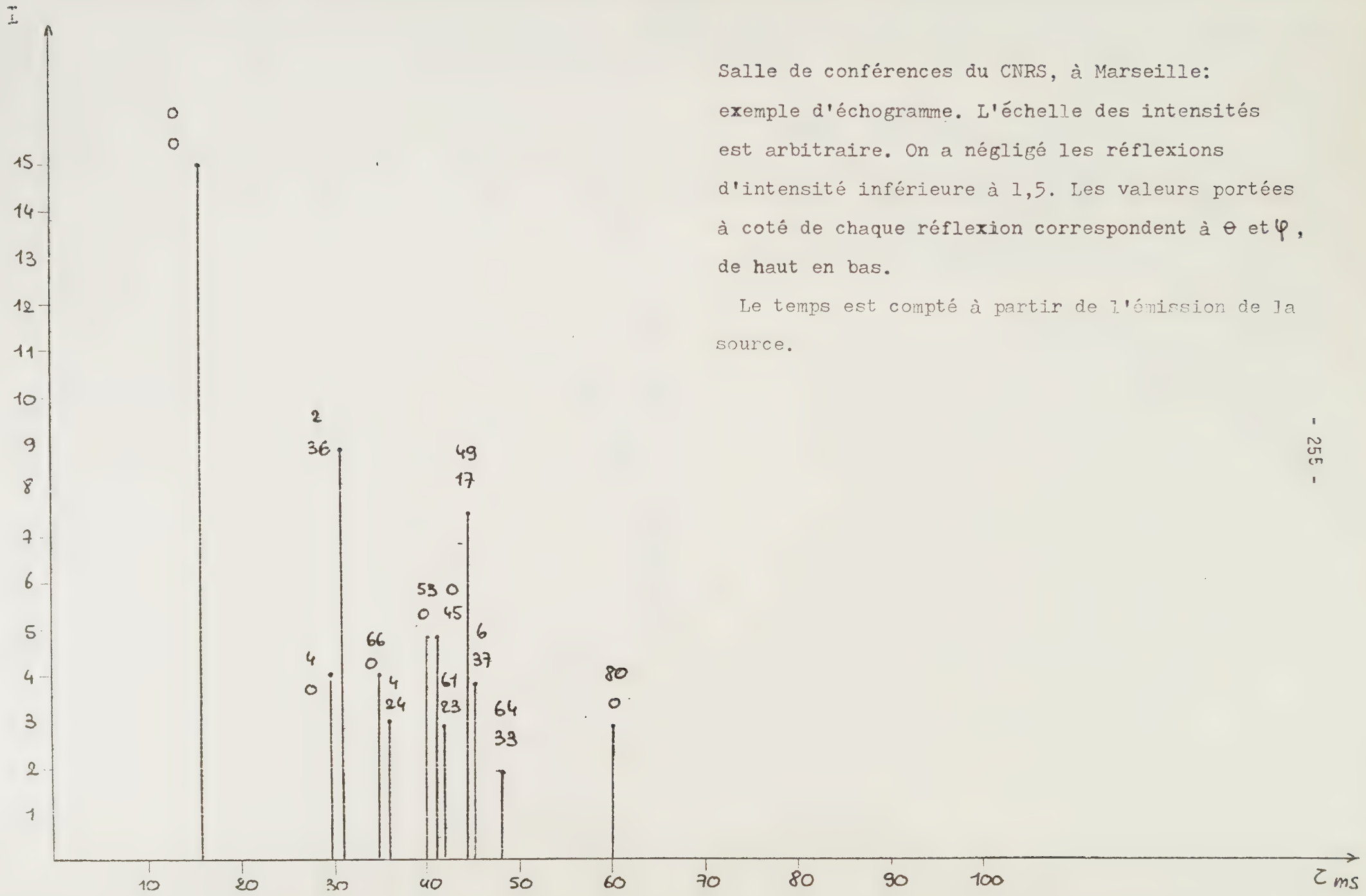
Par contre, si l'on a à envisager une grande salle, on sera obligé d'utiliser les courbes tracées plus haut : il est fort possible que dans ce cas, une estimation globale de l'intelligibilité soit passablement erronée. Toute situation peut alors être considérée comme un "cas particulier". Nous distinguerons deux familles d'échogrammes :

- Pour la première, le son direct est suivi de réflexions proches données par les côtés et le haut d'une "coquille" d'orchestre. Par contre, les réfléxions suivantes risquent d'être assez retardées par rapport au son direct et aux premières réflexions. On est alors dans le cas des figures 5.19 à 5.24.
- La deuxième famille d'échogrammes est celle où le son direct est suivi de réflexions assez éloignées. Dans ce cas, c'est la Fig. 5.10 par exemple qui donnera des renseignements utiles sur le niveau N_{2u} de la première réflexion, d'où l'on déduira à nouveau N_{3u}, et ainsi de suite. Dans l'absolu, ce travail peut paraître inextricable : en fait, dans le cas de grandes salles, les réflexions utiles, ou partiellement utiles, ne seront qu'en nombre limité; au-delà d'un certain délai, de l'ordre de 80 à 100 ms, les réflexions seront toutes nuisibles, constituant un trainage sonore.

Nous espérons ainsi avoir mis en évidence les insuffisantes actuelles des prévisions d'intelligibilité dans les salles, et indiqué comment on peut espérer les améliorer en tenant compte de la répartition spatio-temporelle des réflexions.

==========

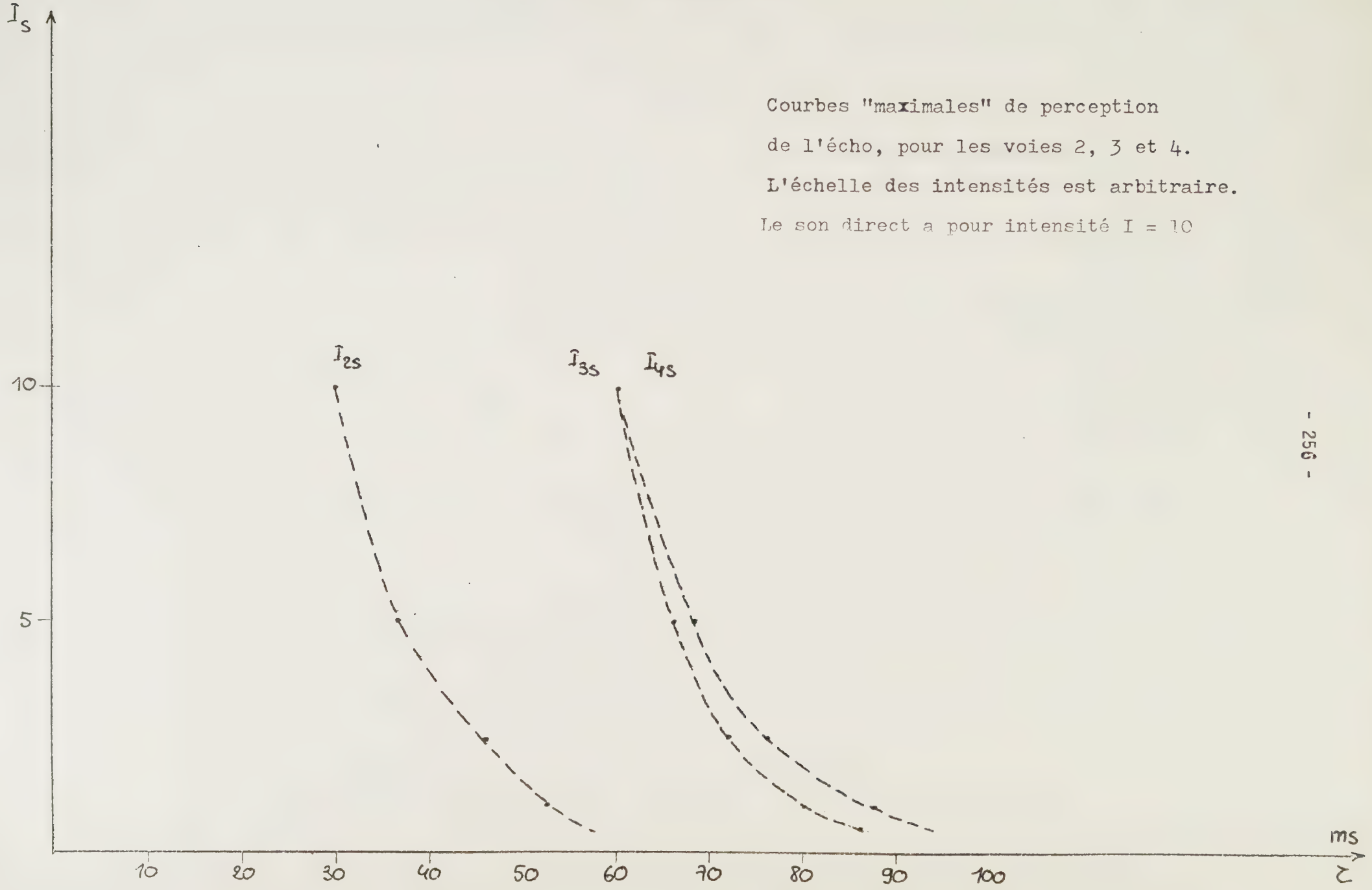
Fig.5.52



Salle de conférences du CNRS, à Marseille: exemple d'échogramme. L'échelle des intensités est arbitraire. On a négligé les réflexions d'intensité inférieure à 1,5. Les valeurs portées à coté de chaque réflexion correspondent à θ et φ, de haut en bas.

Le temps est compté à partir de l'émission de la source.

Fig 5.53



Courbes "maximales" de perception de l'écho, pour les voies 2, 3 et 4. L'échelle des intensités est arbitraire.

Le son direct a pour intensité I = 10

Annexe 1. Calcul du "degré d'écho" de NIESE.

t_{ms}	f(t)	E_s	E	$\sqrt{E}$	$\sqrt{E} - E_s$	S
40	1	1,6	1,7	1,3		0
50	0,85	1,36	4,7	2,17	0,71	0,5
60	0,8	1,26	4,2	2,05	0,77	0,6
70	0,68	1,1	4	2	0,9	0,81
80	0,59	0,93	4,4	2,1	1,17	1,37
90	0,5	0,8	1,2	1,1	0,3	0,09
100	0,41	0,61	3,7	1,92	1,26	1,59
110	0,31	0,5	1,8	1,34	0,84	0,7
120	0,28	0,45	0	0		0
130	0,23	0,37	1,7	1,3	0,97	0,94
140	0,2	0,32	0,9	0,95	0,63	0,4
150	0,16	0,26	1,7	1,3	1	1
160	0,13	0,21	0,6	0,75	0,54	0,29
170	0,11	0,17	0,9	0,95	0,77	0,59
180	0,09	0,15	1,2	1,1	0,95	0,90
190	0,07	0,11	0,7	0,84	0,74	0,55
200	0,01	0,09	0,3	0,55	0,46	0,21
210	0,05	0,08	1,1	1,05	0,98	0,99
220	0,04	0,06	0,4	0,64	0,58	0,34
260	0,02	0,03	0,5	0,72	0,69	0,47
280	,02	0,03	0,2	0,45	0,43	0,19

ANNEXE 2. [illegible] et nuisible,

d'après les [illegible] de LOCHNER et BURGER.

t [illegible]	E [illegible]	[illegible]
0	75	75
10	[illegible]	78
[illegible]	[illegible]	[illegible]45
30	4[illegible]	[illegible]8
40	[illegible]	[illegible]7
50	47	30,5
60	4[illegible]	1[illegible],5
70	41	12,3
80	41	6,15
90	12	1,2
100	37	0

$E_{x} = 132$

$E_{n} = 130$

B I B L I O G R A P H I E

-=-=-=-=-=-=-=-=-=-=-=-=-

5.1 KRYTER (K.).- Method for the calculation and use of the articulation index.- JASA, V 34, n° 11, pp. 1689-97, 1962.-

5.2 KRYTER (K.).- Validation of the articulation index.- JASA, V 34, n° 11, pp. 1698-1702, 1962.-

5.3 BRADY (P.T.).- A technique for investigating on off patterns of speech. The Bell System Technical Journal.- V 44, n° 1, janv. 1965.-

5.4 WEHRHAHN (R.).- Zur Messung des Effektivwertpegels und der Lautstärke von Logatomen.- Tech. Bericht. d. Forschungsinstituts des Fernmeldetechn. , FTZ.A 51 TBr5, juin 1969.- (Cité par KURER [3.42]).-

5.5 HAAS (H.).- The influence of a single echo on the audibility of speech.- J. Audio. Eng. Soc., USA, V 20, n° 2, pp. 146-59, 1972.-

5.6 SERAPHIM (H.P.).- Über die Wahrnehmbarkeit mehrerer Rückwürfe von Sprachschall.- Acustica, V 11, n° 2, pp. 80-91, 1961.-

5.7 SERAPHIM (H.P.).- Raumakustische Nachbildungen mit Elektroakustischen Hilfsmitteln.- Acustica, V 13, pp. 75-85, 1963.-

5.8 MEYER (E.), BURGTORF (W.), DAMASKE (P.).- Eine Apparatur zur Electroakustischen Nachbildung von Schallfeldern.- Acustica, V 15, pp. 339-344, 1965.-

5.9 LEHMANN (P.) et WETTSCHURECK (R.).- Bewertung einer Reflexion in Abhängigkeit von der Richtung.- 7e CIA, Budapest, Communication 24 S 17, 1971.-

5.10 JANUSKA (I.) et SPRONCK (F.).- Essai d'intelligibilité de la parole en milieu bruyant, au moyen de têtes artificielles stéréophoniques.- Electroacoustique (belge), n° 18, pp. 5-37.-

5.11 ANGELINI (Th.) et MEGE (J.).- Enregistreur magnétique à échos.- Note interne CRP, n° 1325, juin 1973.-

5.12 ANGELINI (Th.), MEGE (J.), DAUMAS (A.) et SANTON (F.).- Un enregistreur magnétique à échos.- Revue d'acoustique, n° 28, pp. 19-22, 1974.-

5.13 JANUSKA (I.).- Intelligibilité syllabique de la langue tchèq ue en milieu bruyant.- Symposium "Intelligibilité de la parole", pp. 108-117, Liège, 1973.-

5.14 MUNCEY (R.W.), NICKSON (A.F.B.) et DUBOUT (P.).- The acceptability of speech and music with a single artificial echo.- Acustica, V 3, pp. 168-173, 1953.-

5.15 KARASKIEWICZ (E.) et KWIEK (M.).- Les bases d'une acoustique des salles et le phénomène de Haas.- Acustica, V 13, pp. 25-30, 1963.-

5.16 KARSKIEWICZ (E.) et KWIEK (M.-.- Etude des salles de concert sur des modèles par la méthode des impulsions.- Acustica, V 12, pp. 179-182, 1962.-

5.17 YANAGISAWA (T.) et NIMURA (T.).- Detectable threshold of single echo and sound quality in rooms.- 6e CIA, E-2-5, 1968.-

5.18 PLENGE (G.).- Über die Hörbarkeit kleiner Änderungen der Impulsantwort eines Raumes.- Acustica, V 25, n° 5, pp. 315-325,1971.

5.19 DAMASKE (P.).- Sübjektive Untersuchung von Schallfeldern.- Acustica, V. 19, n° 4, pp. 199-213, 1967/1968.-

5.20 MARSHALL (A.H.).- The importance of room cross section in concert halls.- J. Sound Vibr., V 5, pp. 100-112, 1967.-

5.21 MARSHALL (A.H.).- Levels of reflection masking in concert halls.- J. Sound Vibr., V 7, pp. 116-118, 1968.-

5.22 BARRON (M.).- The subjective effects of first reflections in concert halls and the need of lateral reflections.- J. Sound Vibr., V 15, n° 4, pp. 475-494, 1971.-

5.23 VILLIER KEET (W.).- The influence of early lateral reflections on the spatial impression.- 6e CIA, E-2-4, 1968.-

5.24 KRAUTH (E.) et BÜCKLEIN (R.).- Modellversuche zur Ermittlung des Einflusses von Schallspiegeln auf die Hörsamkeit.- 4e CIA, M 41, 1962.-

5.25 SOMERVILLE (T.), GILFORD (C.), SPRING (N.) et NEGUS (R.).- Recent work on the effects of reflections in concert halls and music studios.- J. Sound Vibr., V 3, pp. 127-134, 1966.-

5.26 MÜLLER (L.).- Zur Klangfärbung durch kurtzeitreflexionen bei Rauschen, Sprache und M usik.- 6e CIA, E-2-6, 1968.-

5.27 MÜLLER (L.).- Zum Mechanismus der klangfärbung in Räumen.- 7e CIA, 24S 19, 1971.-

==========

BIBLIOGRAPHIE SUPPLEMENTAIRE

CHAPITRE 3.-

- Mesures d'intelligibilité :

JANUSKA (I.).- Several statistical problems of subjective intelligibility testing.- 21e Séminaire d'Acoustique, Rzeszow, 1974.-

JANUSKA (I.).- On basic problems of subjective speech intelligibility testing.- 8e Congrès International d'Acoustique, V 1, p. 327, 1974.-

- Estimation de l'intelligibilité de la parole dans les salles:

KURTOVIC (H.).- A contribution to the investigation of useful and disturbing sound reflections.- 8e CIA, V 2, p. 636, 1974.-

BARTHO (R.G.) et HORN (R.P.S.).- An investigation of some computer aided techniques for the evaluation of intelligibility in auditoria.- 8e CIA, V 1, p. 325, 1974.-

CHAPITRE 4.-

- Etude expérimentale d'un échogramme :

LEHMANN (P.).- Zum Einfluss des Sendeortes auf Echogramme und Raumakustische Kriterien.- 8e CIA, V 2, p. 626, 1974.-

NEWMAN (P.I.) et MACKENZIE (R.K.).- Studies of impulse response in model auditoria.- 8e CIA, V 2, p. 606, 1974.-

- Calcul d'un échogramme par une méthode graphique :

STRASZEWICZ (W.) - Méthode d'analyse acoustique spatiale.- 4e CIA, Communication M 26, 1962.-

STRASZEWICZ (W.) - Analyse géométrique des propriétés du champ acoustique dans les espaces clos (en polonais).- Politechnika Warszawska, Prace Naukowe, Elektronika, n° 11, 1974.-

- Calcul d'un échogramme sur ordinateur :

KOLMER (F.) et KYNCL (J.).- Coincidence of reflections. The cause of li-

mited reproducibility of impulse response.- 8e CIA, V2, p. 622, 1974.-

STROM (S.).- Acoustical design of a multi-purpose hall.- 8e CIA, V 2, p. 601, 1974.-

LAMORAL (R.) et SAUVET-GOIGHON (D.).- Etude de grandes salles sur ordinateur.- Acustica, à paraître. 1974.-

OUVRAGE GENERAL.-

KUTTRUFF (H.).- Room Acoustics, Applied Science Publischers, Londres 1973.-

6.- CONCLUSION GENERALE -

Au terme de cette étude, nous résumerons brièvement quelle est la situation actuelle du problème de l'intelligibilité de la parole dans les salles, et comment nous avons pu faire quelque peu avancer cette question.

Le problème de l'intelligibilité est très particulier, et relève de la reconnaissance de l'information sémantique véhiculée par la parole. Or, cette information sémantique dont l'importance, pour l'auditeur, est capitale, ne représente qu'une information assez restreinte de l'information totale de la parole. C'est pourquoi nous n'accordons qu'une confiance limitée aux méthodes qui s'attachent à mesurer la déformation de l'information glabale, et en particulier aux méthodes de corrélation, sans doute mieux adaptées au cas d'une information acoustique de nature essentiellement esthétique, et à la musique en particulier.

Les techniques de prévision de l'intelligibilité les plus fiables, à notre avis, sont celles basées sur les échogrammes. Nous avons tout d'abord entrepris le calcul des échogrammes de façon systématique, en menant de front deux techniques de calcul : celle des images, réservée jusqu'alors au cas simple d'une salle rectangulaire, et celle des rayons. Dans le cas de deux salles réelles, les échogrammes calculés ont été vérifiés expérimentalement, ce qui, à notre connaissance, n'avait jamais été fait.

Les méthodes de prévision de l'intelligibilité basées sur les échogrammes sont des méthodes globales : elles tiennent compte, de façon schématique, du rôle perturbateur apporté par les différents échos. Si ces méthodes donnent de bons résultats, en particulier celle de Januska, il nous a semblé intéressant de savoir si un examen plus rigoureux d'un échogramme ne permettrait pas d'améliorer encore ces prévisions.

La méthode originale que nous proposons est basée sur la mesure du seuil de perception d'un écho du n ième ordre. En effet, alors que les méthodes globales utilisent des mesures d'intégration auditive où un écho est considéré isolèment par rapport au son direct, nous avons évalué

l'énergie nuisible et l'énergie utile d'un écho en fonction de la source primaire constituée par le son direct, et éventuellement les premières réflexions qui lui sont associées. L'influence de la direction a également été prise en considération.

Bien entendu, le sujet n'est pas épuisé : on peut tout d'abord espérer que la physique mathématique nous fournira, dans un avenir proche, des échogrammes calculés sur les bases de l'acoustique ondulatoire, par résolution de l'équation de Dalembert avec des conditions aux limites physiquement réalistes.

D'autre part, on peut envisager de rendre plus souple l'utilisation des courbes du seuil de perception des échos, et de tenir compte éventuellement des modifications apportées par le spectre des réflexions étudiées, dues à des coefficients variables avec la fréquence. Enfin, il ne serait pas sans intérêt de substituer aux mesures données par un seul auditeur expérimenté des statistiques fournies par des groupes d'auditeurs.

Voilà quelles sont les extensions que nous envisageons pour notre travail. Cependant, nous estimons qu'il est d'ores et déjà possible de prévoir sur plans, et sans construction de maquette, quelle sera l'intelligibilité dans une salle à édifier : une application pratique serait à l'heure actuelle particulièrement bienvenue.

-=-=-=-=-=-=-=-=-

Vu, et permis d'imprimer

Marseille, le 21 Novembre 1974

Le Président,

F. BOREL